Amplitude limits and nonlinear damping of shear-Alfvén waves in high-beta low-collisionality plasmas

2017 New J. Phys. 19 055005

(http://iopscience.iop.org/1367-2630/19/5/055005)

New Journal of Physics
The open access journal at the forefront of physics

Deutsche Physikalische Gesellschaft DPG
IOP Institute of Physics

Published in partnership with: Deutsche Physikalische Gesellschaft and the Institute of Physics

OPEN ACCESS

RECEIVED
11 January 2017

REVISED
4 April 2017

ACCEPTED FOR PUBLICATION
6 April 2017

PUBLISHED
17 May 2017

PAPER

Amplitude limits and nonlinear damping of shear-Alfvén waves in high-beta low-collisionality plasmas

J Squire[1,2,6], **A A Schekochihin**[3,4] **and E Quataert**[5]

[1] Theoretical Astrophysics, 350-17, California Institute of Technology, Pasadena, CA 91125, United States of America
[2] Walter Burke Institute for Theoretical Physics, Pasadena, CA 91125, United States of America
[3] The Rudolf Peierls Centre for Theoretical Physics, University of Oxford, 1 Keble Road, Oxford, OX1 3NP, United Kingdom
[4] Merton College, Oxford OX1 4JD, United Kingdom
[5] Astronomy Department and Theoretical Astrophysics Center, University of California, Berkeley, CA 94720, United States of America
[6] Author to whom any correspondence should be addressed.

E-mail: jsquire@caltech.edu

Keywords: Alfvén waves, pressure anisotropy, magnetized turbulence, high-beta plasma, firehose instability, collisionless plasma

Abstract

This work, which extends Squire *et al* (*Astrophys. J. Lett.* 2016 **830** L25), explores the effect of self-generated pressure anisotropy on linearly polarized shear-Alfvén fluctuations in low-collisionality plasmas. Such anisotropies lead to stringent limits on the amplitude of magnetic perturbations in high-β plasmas, above which a fluctuation can destabilize itself through the parallel firehose instability. This causes the wave frequency to approach zero, 'interrupting' the wave and stopping its oscillation. These effects are explored in detail in the collisionless and weakly collisional 'Braginskii' regime, for both standing and traveling waves. The focus is on simplified models in one dimension, on scales much larger than the ion gyroradius. The effect has interesting implications for the physics of magnetized turbulence in the high-β conditions that are prevalent in many astrophysical plasmas.

1. Introduction

In this paper, we derive and discuss stringent nonlinear limits on the amplitude of shear-Alfvén (SA) fluctuations in weakly collisional plasmas. The result, which was first presented in Squire *et al* (2016), is that collisionless linearly polarized SA waves—that is, low-frequency incompressible oscillations of magnetic field ($\delta B_\perp$) and velocity ($u_\perp$) perpendicular to a background field B_0—cannot oscillate when

$$\frac{\delta B_\perp}{B_0} \gtrsim \beta^{-1/2}, \tag{1}$$

where $\beta \equiv 8\pi p_0/B^2$ is the ratio of thermal to magnetic pressure. Above this limit (or a related limit (2) in the weakly collisional regime), standing-wave fluctuations are 'interrupted' before even a quarter oscillation, while traveling waves are heavily nonlinearly damped. In both cases, the magnetic field rapidly forms a sequence of zig-zags—piecewise straight field line segments with zero magnetic tension—and evolves at later times with the magnetic energy far in excess of the kinetic energy (i.e., effectively in a near-force-free state).

What is the cause of such dramatic nonlinear behavior, even in regimes ($\delta B_\perp/B_0 \ll 1$ for $\beta \gg 1$) where linear physics might appear to be applicable? As we now explain, the effect depends on the development of *pressure anisotropy*—i.e., a pressure tensor that differs in the directions perpendicular and parallel to the magnetic field. In a magnetized plasma in which the ion gyro-frequency Ω_i is much larger than the collision frequency ν_c, a decreasing (in time) magnetic field leads to a decreasing pressure perpendicular to the magnetic field ($p_\perp$), while the parallel pressure ($p_\parallel$) increases. Such behavior originates in part from conservation of the particle's first magnetic moment $\mu = mv_\perp^2/2B$, which suggests that $p_\perp/B$ should be conserved as B changes in a collisionless plasma. This anisotropy, $\Delta p \equiv p_\perp - p_\parallel < 0$, provides an additional stress in the momentum equation that can neutralize the restoring effects of magnetic tension, even destabilizing the SA wave and triggering the *parallel firehose instability* if $\Delta p < -B^2/4\pi$ (Rosenbluth 1956, Chandrasekhar *et al* 1958, Parker 1958, Schekochihin *et al* 2010).

Consider the ensuing dynamics if we start with $\Delta p = 0$, but with a field that, in the process of decreasing due to the Lorentz force, *generates a pressure anisotropy that would be sufficient to destabilize the wave*. This is a nonlinear effect not captured in linear models of SA waves. As Δp approaches the firehose limit, the magnetic tension disappears and the Alfvén frequency approaches zero, 'interrupting' the development of the wave. Because the wave perturbs the field magnitude by $\delta B_\perp^2$, an amplitude $\delta B_\perp / B_0 \gtrsim \beta^{-1/2}$ is sufficient to generate such a Δp in a collisionless plasma. As the field decrease is interrupted at the firehose stability boundary, the plasma self-organizes to prevent further changes in field strength, leading to the nullification of the Lorentz force through the development of piecewise-straight (and therefore, tension-less) field-line structures. In addition, as this process proceeds, there is a net transfer of the mechanical energy of the wave to the plasma thermal energy due to 'pressure-anisotropy heating,' which occurs because of spatial correlations between the wave's self-generated pressure anisotropy and $\mathrm{d}B/\mathrm{d}t$.

A similar effect also occurs in the weakly collisional 'Braginskii' regime (Braginskii 1965). Here, collisions act to balance the anisotropy generation and SA waves cannot oscillate if (see equation (1))

$$\frac{\delta B_\perp}{B_0} \gtrsim \sqrt{\frac{\nu_c}{\omega_A}}\,\beta^{-1/2}, \tag{2}$$

where ω_A is the wave frequency and ν_c the ion collision frequency (with $\omega_A \ll \nu_c$ required for the Braginskii equations to be valid). In addition, because a *changing* magnetic field is required to balance the collisional relaxation of Δp, an 'interrupted' wave slowly decays in time until its amplitude is below the limit (2), at which point it can oscillate. Although the details of the nonlinear dynamics differ from the collisionless regime, the dynamics in both regimes share some generic features, in particular the strong dominance of magnetic energy over kinetic energy after a wave is interrupted.

The results described in the previous paragraphs are of interest because the low-frequency SA wave has historically been the most robust plasma oscillation (Cramer 2011). In particular, unlike its cousins, the fast and slow waves, it is linearly unaltered by kinetic physics (except at very high $\beta \gtrsim \Omega_i/\omega_A$; Foote and Kulsrud 1979, Achterberg 1981), and it survives unmodified in even the simplest plasma models (e.g., incompressible magnetohydrodynamics; MHDs). This includes kinetic models of plasma turbulence involving low-frequency, low-amplitude, but fully nonlinear fluctuations (Schekochihin *et al* 2009, Kunz *et al* 2015). For these reasons, SA waves play a key theoretical role in most sub-disciplines and applications of plasma physics: magnetized turbulence phenomenologies (Goldreich and Sridhar 1995, Ng and Bhattacharjee 1996, Boldyrev 2006), the solar wind and its interaction with Earth (Eastwood *et al* 2005, Ofman 2010, Bruno and Carbone 2013), the solar corona (Marsch 2006), solar and stellar interiors (Gizon *et al* 2008), cosmic-ray transport (Schlickeiser 2015), astrophysical disks (Quataert and Gruzinov 1999), and magnetic fusion (Heidbrink 2008), to name a few.

This myriad of applications has in turn led to intense study of the SA wave's basic properties across many plasma regimes (Cramer 2011). The most relevant to our study here are several papers noting that linearly polarized SA waves are Landau damped nonlinearly at the rate $\sim\omega_A \beta^{1/2} (\delta B_\perp/B_0)^2$ at high β (Hollweg 1971b, Lee and Völk 1973, Stoneham 1981, Flå *et al* 1989), although this rate is reduced by particle trapping effects at high wave amplitudes (Kulsrud 1978, Cesarsky and Kulsrud 1981, Völk and Cesarsky 1982). This Landau damping has a similar form to the collisionless 'pressure-anisotropy damping' that plays a key role in some of the effects described in this work. There have also been a wide variety of studies considering nonlinear effects due to parametric instabilities and compressibility (e.g., Galeev and Oraevskii 1963, Hollweg 1971a, Derby 1978, Goldstein 1978, Medvedev and Diamond 1996, Medvedev *et al* 1997, Del Zanna *et al* 2001, Matteini *et al* 2010, Tenerani and Velli 2013), which have generally found large-amplitude SA waves to be unstable to parametric decay at low β, but with stability increasing as β approaches $\sim$1 (Bruno and Carbone 2013). Our study here complements these previous works by showing that in the limit $\beta \gg 1$, linearly polarized finite-amplitude SA waves in weakly collisional plasmas can be nonlinearly modified so strongly that they are unable to oscillate at all. Note, however, that circularly polarized SA fluctuations are unmodified by these effects because their magnetic field strength is constant in time.

The role of SA waves in magnetized turbulence deserves special emphasis: turbulence is fundamental to many areas of astrophysics and geophysics and may be significantly modified by the nonlinear amplitude limit. The well-accepted phenomenology of Goldreich and Sridhar (1995) suggests that strong magnetized turbulence should be understood in terms of nonlinear interactions between SA wave packets, which cascade in such a way that their linear physics is of comparable importance to their nonlinear interactions (this is known as 'critical balance'). Because of the resilience of SA waves to kinetic physics, it is often assumed—and patently true in some cases, e.g., the solar wind at $\beta \lesssim 1$—that Alfvénic cascades survive in collisionless plasmas (Schekochihin *et al* 2009) even though naive estimates suggest the plasma viscosity is very large[7]. The nonlinear interruption of

[7] As recently argued by Verscharen *et al* (2016) for the solar wind, large-amplitude compressive fluctuations may also play an important role in high-β turbulence, aiding in the isotropization of the distribution function.

Alfvénic fluctuations above the amplitude $\delta B_\perp/B_0 \sim \beta^{-1/2}$ may thus significantly alter our understanding of turbulence in weakly collisional plasmas at high β—conditions that occur, for example, in regions of the solar wind (Bale *et al* 2009, Bruno and Carbone 2013), the intracluster medium[8] (Rosin *et al* 2011, Zhuravleva *et al* 2014), and hot astrophysical disks (Balbus and Hawley 1998, Quataert 2001). The picture described above and in what follows suggests a limit on the amplitude (in comparison to a background field) of such turbulence, above which motions are quickly damped, leaving longer-lived magnetic perturbations in their wake.

This paper, which extends the results of Squire *et al* (2016), is organized as follows. In section 2, we present the Landau-fluid (LF) model (Snyder *et al* 1997) used throughout this work to analyze nonlinear SA wave dynamics. This model is chosen as the simplest extension of MHD to weakly collisional plasmas with motions on scales that are large compared to the ion gyroradius. Given the model's relative simplicity in comparison to full Vlasov–Maxwell equations, particular focus is given to gaining qualitative understanding of various physical effects: the pressure anisotropy, collisions, and heat fluxes. Section 3 then contains a very brief description and definition of the two main physical effects—termed *interruption* and *nonlinear damping*—that form the basis for our results. We then treat Braginskii ($\Omega_i \gg \nu_c \gg |\nabla \boldsymbol{u}|$) and collisionless ($\Omega_i \gg |\nabla \boldsymbol{u}| \gg \nu_c$) SA wave dynamics in sections 4 and 5, respectively. Because standing waves are primarily affected by the interruption effect, whereas traveling-wave dynamics are more naturally thought of in terms of nonlinear damping, we split each of these sections and separately discuss standing and traveling waves in each case. For all cases, we derive various scalings, amplitude limits, and damping rates, and describe the physics qualitatively with the aid of numerical examples. In section 6, we discuss the importance of kinetic physics that is not included in our model, both due to the limitations of a 1D domain and of the LF prescription for the heat fluxes. These considerations underscore the importance of future two- and three-dimensional kinetic simulations for further study of the effect. For an impatient reader, the summary of key results in section 7 should be (mostly) comprehensible without reference to the main text.

Finally, the appendices deserve some mention here, being somewhat separate in character and content than the main text of the paper. In these, we derive the nonlinear wave equations asymptotically, both in the collisionless limit (appendix A; we also consider the zero-heat-flux double-adiabatic equations there), and in the Braginskii regime (appendix B). These calculations serve two main purposes. The first is to justify more formally many of the approximations in the main text. In this capacity, they may help comfort a reader who is skeptical of our arguments relating, e.g., to heat fluxes in collisionless waves. The second purpose is to derive explicitly various effects that are only heuristically derived in the main text, e.g., the damping rate for traveling waves. These calculations also provide a useful reference point for future fully kinetic studies that could account more formally for various effects not included in the LF model.

2. Macroscopic equations for a weakly collisional plasma

Throughout this work, our philosophy is to consider the simplest modifications to macroscopic[9] plasma dynamics due to kinetic physics. We thus consider a two-species, fully ionized plasma, and assume that the pressure tensor is *gyrotropic*—i.e., invariant under rotations about the field lines—but can develop an anisotropy, viz., a different pressure parallel and perpendicular to the magnetic field lines. This approximation is generally valid for motions on spatiotemporal scales much larger than those relating to ion gyromotion. It leads to the following equations for the magnetic field and the first three moments of the plasma distribution function (Chew *et al* 1956, Kulsrud 1983, Schekochihin *et al* 2010):

$$\partial_t \rho + \nabla \cdot (\rho \boldsymbol{u}) = 0, \tag{3}$$

$$\rho(\partial_t \boldsymbol{u} + \boldsymbol{u} \cdot \nabla \boldsymbol{u}) = -\nabla\left(p_\perp + \frac{B^2}{8\pi}\right) + \nabla \cdot \left[\hat{\boldsymbol{b}}\hat{\boldsymbol{b}}\left(\Delta p + \frac{B^2}{4\pi}\right)\right], \tag{4}$$

$$\partial_t \boldsymbol{B} = \nabla \times (\boldsymbol{u} \times \boldsymbol{B}), \tag{5}$$

$$\partial_t p_\perp + \nabla \cdot (p_\perp \boldsymbol{u}) + p_\perp \nabla \cdot \boldsymbol{u} + \nabla \cdot (q_\perp \hat{\boldsymbol{b}}) + q_\perp \nabla \cdot \hat{\boldsymbol{b}} = p_\perp \hat{\boldsymbol{b}}\hat{\boldsymbol{b}} : \nabla \boldsymbol{u} - \nu_c \Delta p, \tag{6}$$

$$\partial_t p_\parallel + \nabla \cdot (p_\parallel \boldsymbol{u}) + \nabla \cdot (q_\parallel \hat{\boldsymbol{b}}) - 2q_\perp \nabla \cdot \hat{\boldsymbol{b}} = -2p_\parallel \hat{\boldsymbol{b}}\hat{\boldsymbol{b}} : \nabla \boldsymbol{u} + 2\nu_c \Delta p. \tag{7}$$

Here Gauss units are used, $\boldsymbol{u}$ and $\boldsymbol{B}$ are the ion flow velocity and magnetic field, $B \equiv |\boldsymbol{B}|$ and $\hat{\boldsymbol{b}} = \boldsymbol{B}/B$ denote the field strength and direction, ρ is the mass density, ν_c is the ion collision frequency, $p_\perp$ and $p_\parallel$ are the components of the pressure tensor perpendicular and parallel to the magnetic field, and $q_\perp$ and $q_\parallel$ are fluxes of perpendicular and parallel heat in the direction parallel to the magnetic field. Note that $p_\perp$ and $p_\parallel$ in equation (4)

[8] This is the hot plasma that fills the space between galaxies in clusters.

[9] Here 'macroscopic' refers to scales that are large compared to the plasma microscales, i.e., to those scales that relate to the gyrofrequency, particle Larmor radius, plasma frequency, and Debye length.

are summed over both particle species, while ρ and $\boldsymbol{u}$ in equations (3)–(5) are the ion density and flow velocity (although for $k\rho_i \ll 1$ and $m_e/m_i \ll 1$, they may equivalently be viewed as the total density and flow velocity). The pressure equations (6) and (7) should in principle be solved separately for each species; however, in this work we consider only the ion pressures, an approximation that may be formally justified by an expansion in the electron–ion mass ratio when the electrons are moderately collisional (see, e.g., appendix A of Rosin *et al* 2011). The double-dot notation used in equations (6) and (7) means $\hat{\boldsymbol{b}}\hat{\boldsymbol{b}} : \nabla \boldsymbol{u} \equiv \hat{b}_i \hat{b}_j \nabla_i u_j = \hat{\boldsymbol{b}} \cdot (\hat{\boldsymbol{b}} \cdot \nabla \boldsymbol{u})$. Note that nonideal corrections to the magnetic-field evolution are not included in equation (5) and will be ignored throughout this work (see, e.g., Schekochihin *et al* 2010). These can be important for motions at scales approaching ρ_i and include the Hall effect, electron inertia terms, and resistivity from electron–ion collisions. We also define $\Delta \equiv \Delta p/p_0$ with $p_0 = 2p_\perp/3 + p_\parallel/3$ (note that $\Delta p \ll p_0$ for $\beta = 8\pi p_0/B^2 \gg 1$), the Alfvén speed $v_A = B_0/\sqrt{4\pi\rho}$ (with B_0 a constant background field), the sound speed $c_s = \sqrt{p_0/\rho}$, parallel sound speed $c_{s\parallel} = \sqrt{p_\parallel/\rho}$, and denote the ion gyroradius and gyrofrequency ρ_i and Ω_i, respectively. Although equations (3)–(7) are derived directly from the Vlasov equation assuming $k\rho_i \ll 1$ and $\omega/\Omega_i \ll 1$ (where k and ω are characteristic wavenumbers and frequencies of the system), the heat fluxes $q_{\perp,\parallel}$ remain unspecified and must be solved for using some closure scheme (or the full kinetic equation) as discussed below.

2.1. The importance of pressure anisotropy at high β

In a changing magnetic field, the terms

$$\hat{\boldsymbol{b}}\hat{\boldsymbol{b}} : \nabla \boldsymbol{u} = \frac{1}{B}\frac{\mathrm{d}B}{\mathrm{d}t} + \nabla \cdot \boldsymbol{u} \tag{8}$$

(where $\mathrm{d}/\mathrm{d}t$ is the Lagrangian derivative) in equations (6) and (7) locally force a pressure anisotropy $\Delta = \Delta p/p_0$. Importantly, because this anisotropy generation depends on $\hat{\boldsymbol{b}}$ rather than $\boldsymbol{B}$, its dynamical influence increases as β increases (aside from the limiting effects of microsinstabilities; see below). Namely, the final term of equation (4) shows that Δp has a strong dynamical influence (i.e., is comparable to the Lorentz force) when $\Delta p \sim B^2$; i.e., when $\Delta \sim \beta^{-1}$. For $\beta > 1$, the pressure anisotropy generated by changing B will generally cause a stress that is *stronger* than the Lorentz force. It is also worth noting the importance of the spatial form of Δp, which, as we shall show, can strongly influence the nonlinear dynamics. As will become clear below (sections 2.1.1 and 2.1.2), this spatial variation in Δp depends on the balance between the driving $\hat{\boldsymbol{b}}\hat{\boldsymbol{b}} : \nabla \boldsymbol{u}$ and the other terms in equations (6) and (7) (e.g., the heat fluxes or collisionality), so we should expect nonlinear wave dynamics to depend significantly on a particular physical regime.

In this work, we focus on two such regimes for the evolution of Δp, neglecting compressibility for simplicity in both cases (this neglect is valid at $\beta \gg 1$, $\delta B_\perp/B_0 \ll 1$; see A.3 around equation (A.34) and appendix B around equation (B.7)). The first approximation is Braginskii MHD, which is valid in weakly collisional plasmas when $\Omega_i \gg \nu_c \gg |\nabla \boldsymbol{u}|$; the second is collisionless ($\nu_c = 0$, or equivalently $\Omega_i \gg |\nabla \boldsymbol{u}| \gg \nu_c$), which we model using a simple LF closure for the heat flux.

2.1.1. Braginskii MHD

When collisions dominate ($|\nabla \boldsymbol{u}| \ll \nu_c$), we may neglect $\partial_t p_\perp$ and $\partial_t p_\parallel$ in comparison to $\nu_c \Delta p$ in equations (6) and (7). For $\beta \gtrsim 1$, these approximations also imply $\Delta p \ll p_0$, leading to

$$\Delta p \approx \frac{p_0}{\nu_c}\left(\hat{\boldsymbol{b}}\hat{\boldsymbol{b}} : \nabla \boldsymbol{u} - \frac{1}{3}\nabla \cdot \boldsymbol{u}\right) \approx \frac{p_0}{\nu_c}\frac{1}{B}\frac{\mathrm{d}B}{\mathrm{d}t}. \tag{9}$$

We have neglected $q_{\perp,\parallel}$ for simplicity in deriving equation (9), although this is only valid in the limit $\delta p_{\perp,\parallel}/p_{\perp,\parallel} \ll |\boldsymbol{u}|/c_s$ (where $\delta p_{\perp,\parallel}$ denotes the spatial variation in $p_{\perp,\parallel}$; see Mikhailovskii and Tsypin 1971, Rosin *et al* 2011)[10]. An expression for Δp with heat fluxes included is derived in appendix B (equations (B.11) and (B.12)), where we also briefly discuss how the nonlinear SA wave dynamics are modified by the resulting different spatial form of Δp. However, given the extra complexity of including this effect, we ignore the heat fluxes in the discussion of Braginskii dynamics in section 4.

2.1.2. Collisionless plasma

The evolution of Δ is strongly influenced by heat fluxes when $\nu_c \ll |\nabla \boldsymbol{u}|$ and $\beta \gtrsim 1$. As a simple prescription, we employ a LF closure (Hammett and Perkins 1990, Snyder *et al* 1997, Hammett *et al* 1992, Passot *et al* 2012), which has been extensively used in the fusion community, and to a lesser degree for astrophysical applications (Sharma *et al* 2006, 2007). The heat fluxes are chosen to reproduce linear Landau damping rates, namely,

[10] For $\Delta p \ll p_0$, this condition is approximately equivalent to $\nu_c \sim c_s/\lambda_{\mathrm{mfp}} \gg k_\parallel c_s$ (where λ_{mfp} is the ion mean-free path).

$$q_\perp = -\frac{2c_{s\parallel}^2}{\sqrt{2\pi}\,c_{s\parallel}|k_\parallel| + \nu_c}\left[\rho\nabla_\parallel\left(\frac{p_\perp}{\rho}\right) - p_\perp\left(1 - \frac{p_\perp}{p_\parallel}\right)\frac{\nabla_\parallel B}{B}\right], \tag{10}$$

$$q_\parallel = -\rho\frac{8c_{s\parallel}^2}{\sqrt{8\pi}\,c_{s\parallel}|k_\parallel| + (3\pi - 8)\nu_c}\nabla_\parallel\left(\frac{p_\parallel}{\rho}\right), \tag{11}$$

where $\nabla_\parallel$ is the parallel gradient operator, and the parallel wavenumber $|k_\parallel|$ must be considered as an operator[11]. In the regime of interest, $\Delta p \ll p_0$ and $\nu_c = 0$, with small perturbations to the magnetic field, the dynamical effect of $q_{\perp,\parallel}$ can be easily understood. Equations (10) and (11) are

$$q_\parallel \approx -\sqrt{\frac{8}{\pi}}\rho c_s\frac{\nabla_\parallel}{|k_\parallel|}\left(\frac{p_\parallel}{\rho}\right),\ q_\perp \approx -\sqrt{\frac{2}{\pi}}\rho c_s\frac{\nabla_\parallel}{|k_\parallel|}\left(\frac{p_\perp}{\rho}\right). \tag{12}$$

These, combined with $\hat{\boldsymbol{b}}\cdot\nabla q_{\perp,\parallel} \gg q_{\perp,\parallel}\nabla\cdot\hat{\boldsymbol{b}}$ (valid for small perturbations to the background field), imply that the heat-flux contributions to the pressure equations (6) and (7) simplify to

$$\partial_t p_\perp \sim -\rho c_s|k_\parallel|(p_\perp/\rho),\quad \partial_t p_\parallel \sim -\rho c_s|k_\parallel|(p_\parallel/\rho). \tag{13}$$

These terms, which model the Landau damping of temperature perturbations, suppress spatial variation in $p_{\perp,\parallel}$ over the particle crossing time[12] $\tau_{\rm damp} \sim (|k_\parallel|\,c_s)^{-1}$. This damping implies that if $\tau_{\rm damp} \ll |\nabla\boldsymbol{u}|^{-1} \sim \omega_A$, the $k_\parallel \neq 0$ part of Δ is suppressed by a factor of $\sim\nu_A/c_s \sim \beta^{-1/2}$ compared to its mean[13]. This leads us to the simple interpretation that the heat fluxes *spatially average* Δp, by damping $k_\parallel \neq 0$ components of the pressure perturbations, giving

$$\Delta = 3\int\langle\hat{\boldsymbol{b}}\hat{\boldsymbol{b}} : \nabla\boldsymbol{u}\rangle\,{\rm d}t\,[1 + \mathcal{O}(\beta^{-1/2})(\boldsymbol{x})] \approx 3\left\langle\ln\frac{B(t)}{B(0)}\right\rangle, \tag{14}$$

where $\langle\cdot\rangle$ denotes the spatial average. The spatial form of the $\mathcal{O}(\beta^{-1/2})(\boldsymbol{x})$ term generally follows the spatial variation of $\hat{\boldsymbol{b}}\hat{\boldsymbol{b}} : \nabla\boldsymbol{u}$, and is calculated by asymptotic expansion in various regimes in appendixes A.3 and A.4 (see equations (A.43) and (A.59); these calculations also justify more formally the spatial-averaging action of the heat fluxes derived heuristically above).

2.2. Microinstabilities

An important limitation of equations (3)–(7), which exists for both the Braginskii and LF closures, is their inability to capture correctly certain plasma microinstabilities. For our purposes, at high β, the most important of these are the firehose and mirror instabilities. Both of these grow fastest on scales approaching the Larmor radius, which are explicitly outside the validity of equations (3)–(7). Assuming $\Delta p \ll p_0$, the firehose is unstable if $\Delta \lesssim -2/\beta$ and comes in two flavors: the parallel firehose, which is present in fluid models and is the cause of SA wave interruption, and the oblique firehose (Yoon *et al* 1993, Hellinger and Matsumoto 2000), which grows fastest at $k_\perp \neq 0$, and is not correctly captured by equations (3)–(7). The mirror instability is unstable if $\Delta \gtrsim 1/\beta$ and grows with $k_\perp \gg k_\parallel$. Although the linear mirror instability is contained in the LF model (Snyder *et al* 1997), its nonlinear evolution, which involves trapped particle dynamics (Rincon *et al* 2015), presumably requires a fully kinetic model. It is worth noting that 1D fully kinetic simulations would also not correctly include either the oblique firehose or mirror instabilities.

It has been common in previous literature (e.g., Sharma *et al* 2006, Kunz *et al* 2012, Santos-Lima *et al* 2014) to model the effect of these instabilities in fluid simulations by applying 'hard-wall' boundaries on Δ, limiting its value by the appropriate microinstability threshold. This is motivated by the fact that both in kinetic simulations and, it appears, in the observed solar-wind, microinstabilities act to limit the pressure anisotropy at its marginal values (see, for example, Hellinger *et al* 2006, Bale *et al* 2009, Kunz *et al* 2014, Servidio *et al* 2015). In addition, the enormous scale separation between the micro- and macroscales in many astrophysical plasmas implies that the effect of microscale instabilities on large-scale dynamics should be effectively instantaneous (Melville *et al* 2016). Motivated by the fact that the parallel firehose instability *is* contained in fluid models, most of the numerical

[11] There is some ambiguity in equations (10) and (11) when $c_{s\parallel}$ varies in space, and various more complicated forms may derived when this is an issue (e.g., equation (52) of Snyder *et al* 1997). Because we consider the high-β regime, spatial variation in $c_{s\parallel}$ is very small (see, e.g., appendix A.3.1), so we simply take $c_{s\parallel}$ to be constant for the $2c_{s\parallel}^2(\sqrt{2\pi}\,c_{s\parallel}|k_\parallel| + \nu_c)^{-1}$ and $8c_{s\parallel}^2[\sqrt{8\pi}\,c_{s\parallel}|k_\parallel| + (3\pi - 8)\nu_c]^{-1}$ operators in equations (10) and (11).

[12] Because the spatial variation in ρ will generally be similar to that of $p_{\perp,\parallel}$, the effective damping is less than what it would be if the variation in ρ were ignored in equation (13). However, because the spatial variation in $T_{\perp,\parallel}$ is of the same order as that of $p_{\perp,\parallel}$, a damping of the pressure alone $-c_s|k_\parallel|p_\perp$ may be used for heuristic estimates. A full asymptotic calculation of the relative contributions of ρ and $p_{\perp,\parallel}$ is given in appendix A (see equations (A.41), (A.42) and (A.57), (A.58)).

[13] This estimate arises from the balance between the driving, on timescale $\omega_A^{-1} \sim (k_\parallel\nu_A)^{-1}$, and the damping, on timescale $\tau_{\rm damp} \sim (|k_\parallel|\,c_s)^{-1}$. It is derived in detail in appendix A; see equation (A.43).

results in this work will (where appropriate) apply a limit on positive anisotropies (to model the action of the mirror instability), but not on negative anisotropies.

A more thorough discussion of microinstabilities is given in section 6, focusing in particular on the implications of previous kinetic results for SA wave dynamics and the possible changes that might result from a multi-dimensional fully kinetic treatment.

2.3. Energy conservation

Energy conservation arguments are used heavily throughout the paper, forming the basis for our estimates of traveling-wave damping rates in sections 4.2 and 5.2. With the kinetic, magnetic, and thermal energies defined as

$$E_K = \int \mathrm{d}\boldsymbol{x} \frac{\rho u^2}{2}, \quad E_M = \int \mathrm{d}\boldsymbol{x} \frac{B^2}{8\pi} \quad E_{\mathrm{th}} = \int \mathrm{d}\boldsymbol{x} \left(p_\perp + \frac{p_\parallel}{2} \right), \tag{15}$$

Equations (3)–(7) conserve the total energy:

$$\partial_t (E_K + E_M + E_{\mathrm{th}}) = 0. \tag{16}$$

A key difference compared to standard MHD arises in the evolution equation for the mechanical energy $E_{\mathrm{mech}} = E_K + E_M$:

$$\partial_t E_{\mathrm{mech}} = -\partial_t E_{\mathrm{th}} = \partial_t (E_K + E_M) = \int \mathrm{d}\boldsymbol{x}\, p_\parallel \nabla \cdot \boldsymbol{u} - \int \mathrm{d}\boldsymbol{x}\, \Delta p \frac{1}{B} \frac{\mathrm{d}B}{\mathrm{d}t}. \tag{17}$$

The final term in this equation describes the transfer of mechanical to thermal energy due to the presence of a spatial correlation between Δp and $B^{-1}\mathrm{d}B/\mathrm{d}t$. In the (high-$\beta$) Braginskii limit, where $\Delta p \propto B^{-1}\mathrm{d}B/\mathrm{d}t$ (see equation (9)), this term is always positive and represents a parallel viscous heating. In the collisionless case, it can in principle have either sign, although we shall see that for SA waves, there is a positive correlation between Δp and $B^{-1}\mathrm{d}B/\mathrm{d}t$ that leads to net damping of the waves.

For later reference, the mean pressure anisotropy evolves according to

$$\partial_t \int \mathrm{d}\boldsymbol{x}\, \Delta p = 2 \int \mathrm{d}\boldsymbol{x}\, p_\parallel \nabla \cdot \boldsymbol{u} - 3 \int \mathrm{d}\boldsymbol{x}\, q_\perp \nabla \cdot \hat{\boldsymbol{b}} + \int \mathrm{d}\boldsymbol{x} (p_\perp + 2p_\parallel) \frac{1}{B} \frac{\mathrm{d}B}{\mathrm{d}t} - 3\nu_c \int \mathrm{d}\boldsymbol{x}\, \Delta p. \tag{18}$$

2.4. SA wave dynamics

It is helpful to derive a simple wave equation that isolates the key features of linearly polarized SA waves and the influence of the pressure anisotropy. Although here the derivation is heuristic, with the aim of highlighting the key features of high-β SA dynamics, similar equations are derived asymptotically from the full LF system (3)–(7) in the appendices, for a variety of different regimes (see equations (A.28), (A.38), (A.60), and (B.15)).

Our geometry is that of a background field $B_0\hat{\boldsymbol{z}}$, with perturbations perpendicular to $\hat{\boldsymbol{z}}$ and the wavevector $\boldsymbol{k} = k_z\hat{\boldsymbol{z}} + \boldsymbol{k}_\perp$. For simplicity, we assume the background plasma to be homogenous and stationary (i.e., there is no background flow velocity). Because SA waves are unmodified by $\boldsymbol{k}_\perp \neq 0$ (the envelope is simply modulated in the perpendicular direction) and we analyze only linear polarizations, we assume x-directed perturbations that depend only on z and t, viz.,

$$\boldsymbol{B} = B_0\, \hat{\boldsymbol{z}} + \delta B_\perp(z,\, t)\hat{\boldsymbol{x}}, \quad \boldsymbol{u} = \delta u_\perp(z,\, t)\hat{\boldsymbol{x}}. \tag{19}$$

Note that circularly polarized fluctuations are unaffected by the pressure-anisotropic physics because the field strength remains constant in time. Combining equations (4) and (5) and neglecting compressibility, the field perturbation $\delta b = \delta B_\perp / B_0$ satisfies

$$\frac{\partial^2}{\partial t^2} \delta b = v_A^2 \frac{\partial^2}{\partial z^2} \left[\delta b + \frac{\delta b}{1 + \delta b^2} \frac{\beta \Delta(z)}{2} \right], \tag{20}$$

where Δ is given by equation (14) (collisionless closure) or equation (9) (Braginskii closure). In the absence of a background pressure anisotropy, equation (20) illustrates that linear long-wavelength SA fluctuations are unmodified by kinetic effects. Similarly, fixing Δ and linearizing in δb, the parallel firehose instability emerges because the coefficient of $\partial_z^2 \delta b$ is negative for $\beta\Delta/2 < -1$.

In the following sections, we shall treat standing and traveling waves separately. While these differ only in their initial conditions, they can display rather different nonlinear dynamics. In the context of equation (20), a standing wave has initial conditions in either δb or u_x/v_A, viz.,

$$\delta b(t = 0) = -\delta b_0 \cos(k_z z),\ u_x(t = 0) = 0, \tag{21}$$

or

$$\delta b(t = 0) = 0,\ u_x(t = 0)/v_A = \delta b_0 \sin(k_z z); \tag{22}$$

a traveling wave involves initial conditions in both δb and u_x/v_A, viz.,

$$\delta b\,(t=0) = -\delta b_0 \cos(k_z z),\ u_x(t=0)/v_A = \delta b_0 \cos(k_z z) \tag{23}$$

(for a wave traveling from left to right).

2.5. Numerical method

For all numerical examples, both of the LF equations and of various reduced equations (in the appendices), we use a simple Fourier pseudospectral numerical method on a periodic domain. Standard 3/2 dealiasing is used, along with a k^6 hyperviscous diffusion operator in all variables, which is tuned so as to damp fluctuations at scales just above the grid scale. This is necessary with Fourier methods because there is little energy dissipation otherwise, and the energy can be spuriously reflected back from high-k into lower-k modes. The only further approximation used in solving equations (3)–(11) is the identification of $|k_\parallel|$ in equations (10) and (11) with $|k_z|$ (the $|k_\parallel|$ operator is nondiagonal in both Fourier and real space and thus somewhat complex to evaluate). While this approximation is truly valid only for $\delta b \ll 1$, various tests have shown that the exact form of the heat fluxes makes little difference; for example, the method of Sharma *et al* (2006), which sets $|k_\parallel| = k_L$ with k_L a parameter, does not qualitatively modify the solutions presented here. Results shown in the figures throughout the text were obtained at a resolution $N_z = 512$, but we see little modification of results at higher or lower resolutions.

3. Wave interruption and damping through pressure anisotropy

In this section, we explain the two key mechanisms that can lead to strongly nonlinear behavior of SA waves in high-β regimes. These are: (1) the nullification of the wave's restoring force (the Lorentz force) through the self-generated pressure anisotropy, which we term *interruption*; and (2) the channeling of wave energy into thermal energy due to spatial correlation of Δ and $\mathrm{d}B/\mathrm{d}t$, which we term *nonlinear damping*.

3.1. Interruption

It is immediately clear from equation (20) that any time $\Delta(z)$ approaches $-2/\beta$, the solutions to equation (20) are fundamentally altered because the restoring force of the SA wave disappears (i.e., the coefficient of ∂_z^2 approaches zero). We term this effect 'wave interruption', because the oscillation halts when this occurs.

In the Braginskii limit, with Δ given by equation (9), the wave is thus interrupted when

$$\nu_c^{-1}\frac{1}{B}\frac{\mathrm{d}B}{\mathrm{d}t} \sim -\frac{2}{\beta}. \tag{24}$$

Here $B = B_0(1+\delta b^2)^{1/2}$, which depends only on the current value of the field and its rate of change.

By contrast, in the collisionless limit, with Δ given by equation (14), the interruption occurs if a wave evolves so that

$$3\left\langle \ln\frac{B(t)}{B(0)}\right\rangle = -\frac{2}{\beta}, \tag{25}$$

which is interesting for its explicit dependence on the initial conditions. Our derivations of amplitude limits in the following sections are simply applications of equations (24) and (25).

3.2. Nonlinear damping

In the presence of a positive correlation between Δp and $B^{-1}\mathrm{d}B/\mathrm{d}t$, the mechanical energy of the wave is converted to thermal energy at the rate (see equation (17))

$$\partial_t E_{\mathrm{mech}} = -\int \mathrm{d}\boldsymbol{x}\ \Delta p\frac{1}{B}\frac{\mathrm{d}B}{\mathrm{d}t} = -\partial_t E_{\mathrm{th}}. \tag{26}$$

We term this effect 'nonlinear damping' because the fact that the B perturbation is proportional to δb^2 in a SA wave implies that the damping rate also scales with δb^2.

In the Braginskii limit (equation (9)), $\Delta \propto \nu_c^{-1}B^{-1}\mathrm{d}B/\mathrm{d}t$ and the energy damping rate is, therefore,

$$\partial_t E_{\mathrm{mech}} \sim -\nu_c^{-1}\int \mathrm{d}\boldsymbol{x}\ p_0\left(\frac{1}{B}\frac{\mathrm{d}B}{\mathrm{d}t}\right)^2, \tag{27}$$

which is simply the parallel viscous damping.

In the collisionless case, there is no fundamental requirement that Δp and $B^{-1}\mathrm{d}B/\mathrm{d}t$ have a positive spatial correlation. Nonetheless, given that Δp is driven by $B^{-1}\mathrm{d}B/\mathrm{d}t$ (see equations (6) and (7)), one might intuitively expect such a correlation for SA waves, and the calculations in section 5.2 and appendix A.3 show that this is indeed the case. Note, however, that its numerical value, and thus the wave damping rate, depends on the effect

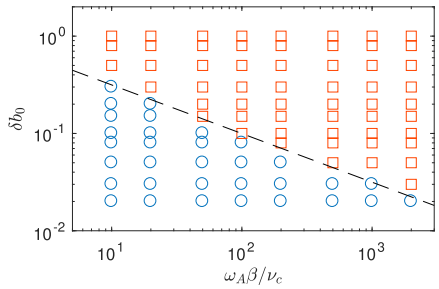

Figure 1. Numerical confirmation of the scaling (30). Each marker represents a numerical solution of the SA wave equation (20) with the Braginskii closure (9), starting from a sinusoidal magnetic perturbation (initial conditions (21)), with amplitude δb_0 and some chosen $\omega_A\beta/\nu_c$ (see figure 2). A red square indicates that an initial perturbation was interrupted before a half cycle (as in figure 2), while a blue circle indicates that the perturbation flipped polarity without interruption. The dashed line is $\delta b_0 = 2.5(\omega_A\beta/\nu_c)^{-1/2}$. Note that in the incompressible limit, SA wave dynamics are determined entirely by δb_0 and the ratio $\omega_A\beta/\nu_c$, because the ν_c^{-1} factor in Δ in equation (9) multiplies $\beta/2$ in equation (20).

of the heat fluxes in smoothing Δp (this is the $\mathcal{O}(\beta^{-1/2})(\boldsymbol{x})$ part in equation (14)). In the collisionless limit, there is also Landau damping of a nonlinear SA wave due to the spatiotemporal variation of the magnetic pressure (Hollweg 1971b, Lee and Völk 1973), which turns out to cause wave damping at a rate similar to the pressure anisotropy damping (neglecting particle trapping effects; Kulsrud 1978).

4. Braginskii MHD—the weakly collisional regime

In this section we work out the behavior of SA waves in the Braginskii limit. As discussed in section 2, Braginskii dynamics differ significantly from fully collisionless dynamics because the pressure anisotropy is determined by the current value of $\partial_t B$, rather than the time history of the magnetic field.

4.1. Standing waves

Starting from a finite-amplitude magnetic perturbation, a Braginskii standing SA fluctuation will be significantly modified (interrupted) if equation (24) is satisfied at some point during its decay. If we consider an unmodified standing wave

$$\delta b(z,\, t) = \delta b_0 \cos(k_z z)\cos(\omega_A t) \tag{28}$$

with $\delta b_0 \ll 1$, then

$$\Delta = \frac{1}{\nu_c}\frac{1}{B}\frac{\mathrm{d}B}{\mathrm{d}t} \approx -\frac{1}{2}\frac{\omega_A}{\nu_c}\delta b_0^2 \sin(2\omega_A t)\cos^2(k_z z). \tag{29}$$

The wave will thus be significantly modified—i.e., interrupted—if $\Delta \lesssim -2/\beta$ at some point in space, which occurs if

$$\delta b_0 \gtrsim 2\sqrt{\frac{\nu_c}{\omega_A}}\,\beta^{-1/2} \equiv \delta b_{\max}. \tag{30}$$

Above this limit, Δp can remove the restoring force of the wave in regions where $\delta b(z) \neq 0$ (i.e., around the antinodes of the wave). As we show in figure 1, the limit (30) is well matched by numerical solutions.

Figure 2 illustrates the dynamics of a standing wave above the limit (30). We solve the SA wave equation (20), using the Braginskii closure (9), which assumes only 1D dynamics and incompressibility of the wave[14]. Note that within the incompressible limit, the dynamics are entirely determined by δb_0 and the ratio $\beta^{-1}\nu_c/\omega_A$, because the ν_c^{-1} factor in Δ (equation (9)) multiplies $\beta/2$ in equation appendix(20).

Although the nonlinear wave dynamics shown in figure 2 may appear quite bizarre, with angular field structures and sharp discontinuities in Δp, many features can be straightforwardly understood by noting that if the field is to decrease significantly more slowly than in a linear SA wave, it must nullify the magnetic tension. This can be achieved: (1) by keeping a pressure anisotropy at the firehose limit, which occurs in the $\delta b(z)$ 'humps' where the field is curved; or (2) by having straight field lines, which occurs where $\delta b(z)$ is zero. Then, because the field must keep decreasing in order to maintain $\Delta = -2/\beta$ (since $\Delta \sim \nu_c^{-1}\mathrm{d}B/\mathrm{d}t$), it slowly decays

[14] Note that for the condition (30) to be met at the same time as the condition $\nu_c \gg \omega_A$, required for the validity of the Braginskii equations, the system must be at very high β. Further, since $u \sim v_A \sim \beta^{-1/2}c_s$, the motions are very subsonic. It thus makes sense to assume incompressibility when studying Braginskii waves, and if one wished to study the lower-β, larger-δb_0, limit between the collisionless and Braginskii regimes, it would be most sensible to solve the full LF equations (3)–(7) with the collisional relaxation terms included. More discussion, including the effects of heat fluxes, is given in appendix B.1.

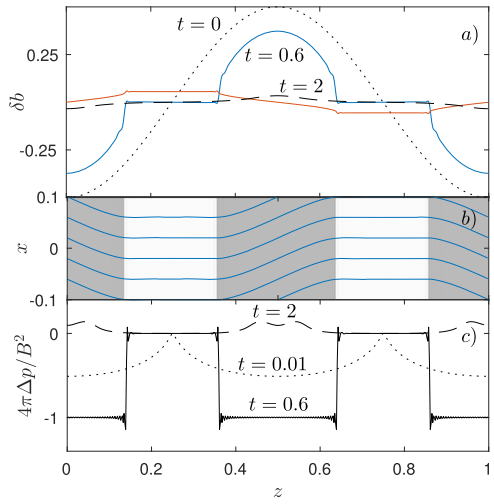

Figure 2. Evolution of an initial magnetic perturbation $\delta b = -0.5\cos(2\pi z)$ within the Braginskii model at $\sqrt{\nu_c/\omega_A}\beta^{-1/2} = 0.2$. Panel (a) shows δb at $t = 0$ (black dotted line), δb at $t = 0.6\tau_A$ (blue solid line), $u_\perp/v_A$ at $t = 0.6\tau_A$ (red solid line), and δb at $t = 2\tau_A$ (black dashed line), which is after the amplitude has decreased below the interruption limit. Panel (b) illustrates the shape of the magnetic field lines in space at $t = 0.6\tau_A$ (blue lines), with the shading showing where $4\pi\Delta p/B^2 = 0$ (white) or -1 (gray). Panel (c) shows the anisotropy parameter, $4\pi\Delta p/B^2$, which is -1 at the parallel firehose limit, at the same times as in panel (a) (the black dotted line shows $t = 0.01\tau_A$ to illustrate the initial evolution). Note that the velocity peturbation is much smaller than the magnetic peturbation during the decay, and the lack of magnetic tension everywhere in the decaying wave.

in time, with the regions where the field is small reaching $\delta b(z) = 0$ first. Once the amplitude of the wave decays below the level at which it can sustain $\Delta = -2/\beta$ (i.e., when the field at the wave antinode is $\delta b^2 \sim \nu_c/\omega_A\beta^{-1}$), the wave can oscillate freely again with an amplitude below the interruption limit (although δb is not sinusoidal because the final stages of the interrupted decay are nonsinusoidal). Note that throughout this decay process, the perturbation's magnetic energy dominates over the kinetic energy. This is because the pressure anisotropy stress cancels out the Lorentz force in the momentum equation, leading to a magnetic field that changes more slowly than that in a similar-amplitude linear SA wave.

We can use these ideas to calculate the decay time of the field, $t_{\rm decay}$, as a function of β and the initial amplitude δb_0. The idea is simply to ignore the spatial dependence of the solution, focusing on the antinode of the wave, where δb is maximal. The condition $\Delta = -2/\beta$ is then

$$-\frac{2}{\beta} = \Delta = \frac{1}{\nu_c}\frac{1}{B}\frac{\mathrm{d}B}{\mathrm{d}t} \sim \frac{1}{\nu_c}\delta b\frac{\partial \delta b}{\partial t}, \tag{31}$$

which has the solution

$$\delta b^2 = \delta b_0^2 - 4\frac{\nu_c}{\beta}t. \tag{32}$$

By solving for $\delta b^2 = 0$, we arrive at a prediction for the time for the interrupted field to decay to an amplitude at which it can oscillate:

$$t_{\rm decay} = \frac{\beta}{4\nu_c}\delta b_0^2. \tag{33}$$

As shown in figure 3, this estimate agrees very well with numerically computed decay times (taken from calculations like that in figure 2) even quantitatively, illustrating the effectiveness of the simple dynamical model proposed above.

4.1.1. Standing waves with an initial velocity perturbation

It is worth briefly describing also the dynamics that one observes after initializing with a velocity rather than a magnetic perturbation[15]. Because in such a situation the field initially grows rather than decays, the ensuing dynamics depend on what occurs at positive pressure anisotropies, when $\Delta > 1/\beta$. Specifically, one expects growing mirror fluctuations (which are not captured in 1D models) to act to limit Δ at $1/\beta$, and that this limiting action will be fast compared to ω_A, so long as there is significant scale separation with the gyroscale (see section 2.2; Kunz *et al* 2014, Melville *et al* 2016). If this limit on Δ does not exist—i.e., if Δ can grow without bound as the magnetic field grows in the standing wave—the extra magnetic tension arising from $\Delta > 0$ acts to

[15] This arguably represents a more natural situation physically, since it is hard to envisage how a static magnetic perturbation might arise.

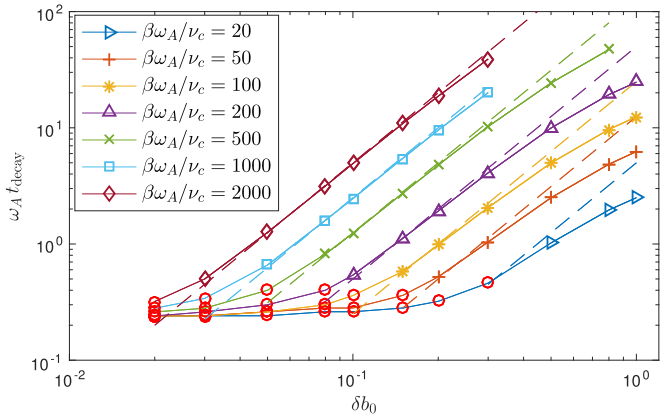

Figure 3. Time $t_{\rm decay}$ for an initial magnetic perturbation of the form $\delta b(z) = \delta b_0 \cos(k_z z)$ to decay to an amplitude that is small enough that it can oscillate. Solid lines and symbols show the numerically measured $t_{\rm decay}$ (normalized by ω_A^{-1}), while dashed lines of matching color show the theoretical prediction (33). The bright red circles are those points for which δb_0 is below the limit (30), meaning the wave is able to oscillate. The match with the theoretical prediction for $t_{\rm decay}$ is surprisingly accurate, illustrating the usefulness of the simple arguments outlined in section 4.1.

reverse the fluctuation of the wave at low magnetic-field amplitudes, while strong nonlinear damping causes the wave to damp to an amplitude below the interruption limit (30) in less than a wave period (see section 4.2). However, if the anisotropy is limited at positive values, this allows the field to grow to much higher amplitudes, viz., $\delta b \approx u_{\perp 0}/v_A$ (for $u_{\perp 0}/v_A \gg \beta^{-1/2}$ and $\beta \gg 1$). When the field then starts decreasing again (once $u_\perp \approx 0$) it does so from an amplitude that is above the interruption limit, and thus behaves in effectively the same way as an initial purely magnetic perturbation. Thus, the dynamics—as long as growing mirror fluctuations act to limit positive pressure anisotropies—are similar to those for an initial static magnetic perturbation, and the limit on the amplitude of an initial velocity perturbation $u_{\perp 0}/v_A$ is similar to equation (30). In section 6, we give a more detailed discussion of this physics[16].

4.2. Traveling waves

With the Braginskii model, because $\Delta \approx \nu_c^{-1} B^{-1} {\rm d}B/{\rm d}t$, the anisotropy and the rate of change of the magnetic field $B^{-1}{\rm d}B/{\rm d}t$ are always strongly correlated. From equation (17), this implies the wave energy $E_{\rm wave} = E_{\rm mech} = E_K + E_M$ is nonlinearly damped at the rate

$$\partial_t E_{\rm wave} = -\frac{1}{\nu_c}\int {\rm d}\boldsymbol{x}\, p_0 \left(\frac{1}{B}\frac{{\rm d}B}{{\rm d}t}\right)^2. \tag{34}$$

A sinusoidal traveling SA wave, $\boldsymbol{B} = B_0(\delta b \sin(k_z z - \omega_A t),\, 0,\, 1)$, creates a changing magnetic field

$$\frac{1}{B}\frac{{\rm d}B}{{\rm d}t} = -\frac{1}{2}\delta b^2 \omega_A \sin(2k_z z - 2\omega_A t), \tag{35}$$

which, from equation (34), causes the wave to damp at the rate

$$\partial_t E_{\rm wave} \approx -\frac{1}{8} p_0 \frac{\omega_A^2}{\nu_c}\delta b^4. \tag{36}$$

This is effectively a parallel viscous damping, which occurs because there is a component of $\boldsymbol{u}$ in the field-parallel direction due to the finite amplitude of the wave. Noting that $E_{\rm wave} = (8\pi)^{-1}\delta b^2 B_0^2$, we conclude that equation (36) implies a wave damping rate

$$\frac{1}{2E_{\rm wave}}\frac{{\rm d}E_{\rm wave}}{{\rm d}t} \approx -\frac{1}{4}\omega_A \frac{\beta\omega_A}{\nu_c}\delta b^2 = -\omega_A \frac{\delta b^2}{\delta b_{\rm max}^2}, \tag{37}$$

where $\delta b_{\rm max}$ is the interruption limit, given by equation (30).

For amplitudes above the interruption limit (30), equation (37) implies a damping rate greater than the frequency of the wave itself. In this case, the damping will cause such a strong nonlinear modification of the wave that it might be considered more accurately as an interruption, effectively stopping the wave. Indeed, the local pressure anisotropy will reach the firehose limit in regions where $B^{-1}{\rm d}B/{\rm d}t < 0$ (and the mirror limit where

[16] The reader may be puzzled that we are not applying such a 'Δ-limit' argument also to the negative anisotropies. The salient point is that Δ can only be limited by firehose microinstabilities once $\Delta \lesssim -2/\beta$, by which point the magnetic tension has already been removed. More extensive discussion of this and related issues is given in section 6.

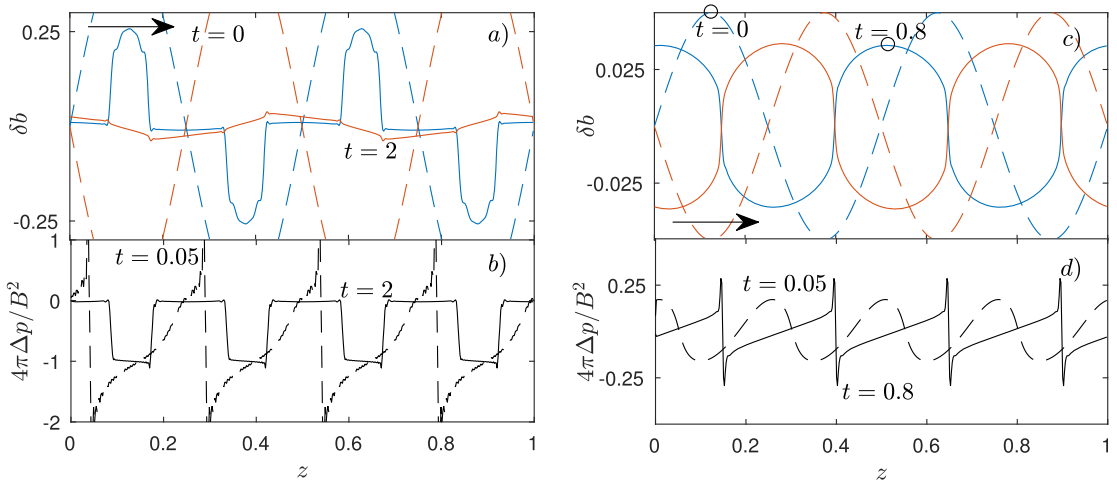

Figure 4. Evolution of a traveling wave with the initial condition $\delta b = -u_\perp/v_A = -\delta b_0 \cos(4\pi z)$ (equation (23)) within the Braginskii model at $\sqrt{\nu_c/\omega_A}\beta^{-1/2} = 0.2$. In the top panels, blue lines illustrate δb and red lines show $u_\perp/v_A$, while the bottom panels show the anisotropy parameter $4\pi\Delta p/B^2$ (this is -1 at the parallel firehose limit). The left panels (a) and (b) show a wave with $\delta b_0 = 0.5$—i.e., above the interruption limit—at $t = 0$ (dashed lines; we show $t = 0.05$ instead for the anisotropy parameter to illustrate the initial evolution) and $t = 2\tau_A$ (solid lines) (the initial condition in panel (a) is sinusoidal; a reduced plot range is used to better see the later times). The right panels (c) and (d) show a wave with $\delta b_0 = 0.05$—i.e., below the interruption limit—at $t = 0$ (dashed lines; $t = 0.05$ in panel (c)), $t = 0.8\tau_A$ (solid lines) (the earlier time is shown so as to fit the full evolution into one panel). In panel (c) hollow circles mark the same position on the wave front as it propagates. Above the interruption limit, there is such a strong nonlinear modification of the wave that it is effectively interrupted and stopped before it can propagate. In contrast, below the limit, the wave undergoes minor shape changes and slow nonlinear decay.

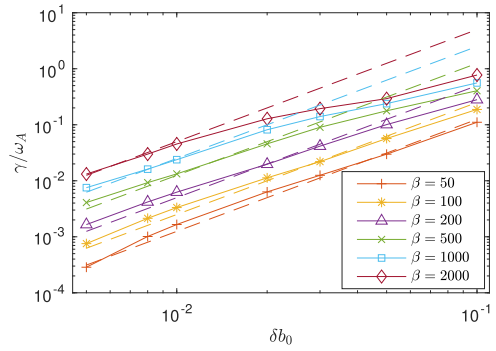

Figure 5. Initial decay rate $\gamma/\omega_A \equiv 0.5\, E_{\rm wave}^{-1}\mathrm{d}E_{\rm wave}/\mathrm{d}t$ of Braginskii traveling waves. Solid lines and symbols show the numerical measurements, while dashed lines are the theoretical prediction (37). The agreement is reasonably good, although at larger decay rates, where the nonlinear effects are stronger, the wave reduces its damping rate by becoming more square.

$B^{-1}\mathrm{d}B/\mathrm{d}t > 0$), so we should expect some of the arguments of section 4.1 to apply here. A traveling wave (with the same parameters as the standing wave in figure 2) is illustrated in figures 4(a) and (b), showing how the wave is virtually stopped with a larger magnetic than kinetic energy, and an anisotropy that is similar to the standing wave (see figure 2(c)). Thus, there is effectively an 'interruption' of the same kind as for a standing wave.

Consider now a traveling wave that is well below the interruption limit, illustrated in figures 4(c) and (d). Such a wave exhibits much slower damping and moderate nonlinear modification to the wave shape, which should be expected because the mechanism causing the wave damping (the pressure anisotropy) has nonlinear spatial variation in space. The more angular structures that the wave develops act to reduce $B^{-1}\mathrm{d}B/\mathrm{d}t$ over much of the wave, and thus reduce the damping rate somewhat.

The damping rate is measured quantitatively in figure 5, where the theoretical prediction (37) is compared to the rates measured in simulations. While our prediction agrees very well in the low-decay-rate limit, where the nonlinear modifications to the wave shape are small, it deviates as the wave damping increases and the pressure anisotropy causes more significant changes to the shape of the wave.

Finally, note that although, for consistency with the upcoming analysis of collisionless plasmas (section 5), we have discussed standing waves and traveling waves separately, this distinction is less important for Braginskii dynamics. Indeed, we have seen that a Braginskii traveling wave above the amplitude limit (30) is effectively

interrupted, because the anisotropy is so strong that it stops the wave. Analogously, a standing wave below the interruption limit will oscillate but will be nonlinearly damped at the rate (37), because there is still parallel variation in $\nabla \boldsymbol{u}$ that is damped by the Braginskii viscosity. In the next section, we shall see that there is a stronger distinction between standing and traveling waves (and between interruption and nonlinear damping) for collisionless wave dynamics, because of the smoothing effect of the heat fluxes.

5. Collisionless waves

We now consider the dynamics of SA waves in a collisionless plasma at high β. Collisionless dynamics differ significantly from the Braginskii limit discussed in the previous section because, in a collisionless plasma, the pressure anisotropy remembers the *time history* of B^2, rather than being set by the instantaneous value of $\mathrm{d}B/\mathrm{d}t$. This implies that, once the magnetic tension is removed when the anisotropy reaches the firehose limit, the field is not able to decrease; any further decrease in B would drive the plasma unstable. In contrast, in the Braginskii limit, maintaining Δp at the firehose limit *requires* $\mathrm{d}B/\mathrm{d}t < 0$. In addition, the heat fluxes always play a significant dynamical role in collisionless plasmas at high β, acting to smooth Δp. This leads to near perfect zig-zag magnetic field lines that minimize the spatial variation in B^2.

Compared to the Braginskii model, which may be rigorously derived from the kinetic equations via a perturbative expansion (Braginskii 1965), LF closures are only heuristically motivated (see section 2). Nonetheless, for the clarity of presentation throughout this section, we shall primarily focus our discussion on physics contained within the LF model, viz., large heat fluxes that result from particles streaming along field lines, with no particle scattering. A variety of other physical effects that may be important (e.g., particle trapping, or particle scattering by magnetic fluctuations due to microinstabilities) are discussed in section 6.

5.1. Standing waves

As discussed in section 2 and more formally justified in appendix A (see equation (A.34) and related discussion), the primary effect of the heat fluxes is to damp all $k_\parallel \neq 0$ components of Δp, giving

$$\Delta = 3\left\langle \ln \frac{B(t)}{B(0)} \right\rangle [1 + \mathcal{O}(\beta^{-1/2})(\boldsymbol{x}) + \cdots]. \tag{38}$$

Since $\langle B(t)/B(0)\rangle$ decreases in time as a standing wave evolves, a wave will reach $\Delta = -2/\beta$ if

$$\frac{3}{2}\left\langle \ln \frac{1}{1+\delta b_0(\boldsymbol{x})^2} \right\rangle \approx -\frac{3}{2}\langle \delta b_0(\boldsymbol{x})^2\rangle < -\frac{2}{\beta}. \tag{39}$$

Assuming a sinusoidal initial perturbation $\delta b_0(\boldsymbol{x}) = \delta b_0 \cos(k_z z)$, a SA wave is interrupted if

$$\delta b_0 \gtrsim \sqrt{\frac{8}{3}}\,\beta^{-1/2} \equiv \delta b_{\max}. \tag{40}$$

This limit, including the $\sqrt{8/3}$ numerical coefficient, matches numerical simulations using the full LF model nearly perfectly (Squire *et al* 2016).

The dynamics of a perturbation that starts *above* the limit (40) are illustrated in figure 6, which shows a solution of the LF equations (3)–(11). Despite the bizarre appearance of the highly angular, zig-zag structures that develop here (see figure 6(c)), the main features can be relatively easily understood. Let us consider qualitatively the wave evolution in three phases:

Approach to interruption. During the initial evolution of the wave, before the anisotropy reaches $\Delta = -2/\beta$, the spatial shape of the wave is largely unaffected by the developing anisotropy. This is because spatial variation of $\beta\Delta$ is $\mathcal{O}(\beta^{-1/2})$ (see equation (38)) even though $\beta\Delta/2$ itself is $\mathcal{O}(1)$ during this phase. The nonlinearity due to the pressure anisotropy $\beta\partial_z^2(\delta b\Delta)$ (see equation (20)) is thus $\sim\beta\Delta\partial_z^2\delta b$, which simply slows down the wave without modifying its spatial structure. The pressure anisotropy during this phase is shown as a dashed line in figure 6(b), while $\delta b(z)$ looks very similar to the initial condition (dotted line in figure 6(b)) at these parameters.

Early nonlinear evolution. As the pressure anisotropy approaches $\Delta = -2/\beta$, the linear term in the wave evolution equation (20), $1 + \beta\Delta/2$, becomes very small and then turns negative when the anisotropy overshoots the parallel firehose limit. This overshoot has two effects: the first is to reverse the decrease in the magnetic field of the largest-scale mode (since the linear term has changed sign); the second is to cause small-scale firehose fluctuations to grow rapidly in the regions of low $\delta b(z)$ (i.e., in the neighborhood of the wave nodes; the overshoot of the firehose limit is greatest at low fields because Δ does not vary significantly in space). These growing small-scale modes act very quickly to return the anisotropy back to its marginal level. This process can be seen in the $t = 0.15\tau_A$ curves in figures 6(a) and (c), which show the field and the pressure anisotropy just

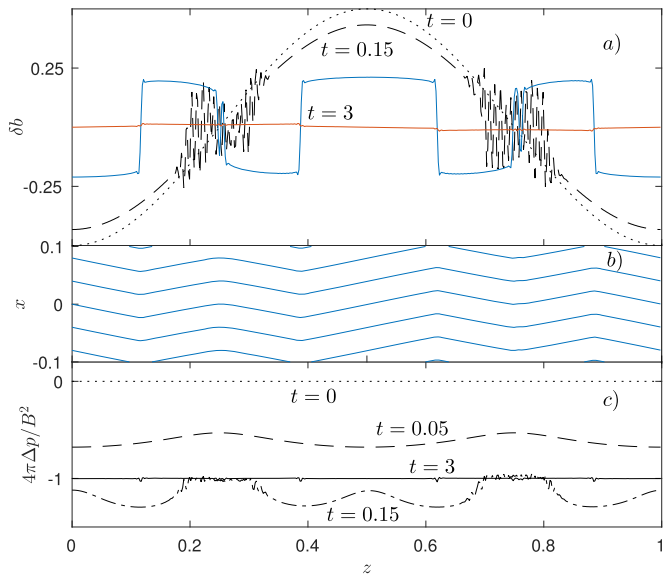

Figure 6. Evolution of a shear-Alfvén standing wave in the collisionless LF model at $\beta = 100$, at an initial amplitude above the interruption limit (40), with initial condition $\delta b = -0.5\cos(k_z z)$. Panel (a) shows δb at $t = 0$, $t = 0.15\tau_A$, and $t = 3\tau_A$ (dotted black, dashed black, and blue lines, respectively), and $u_\perp/v_A$ at $t = 3\tau_A$ (red line). Panel (b) shows the magnetic-field lines of the perturbation shown in (a) at $t = 3\tau_A$. Panel (c) shows the anisotropy $4\pi\Delta p/B^2 = \Delta\beta/2$ associated with the evolution shown in (a), at $t = 0$, $t = 0.05\tau_A$, $t = 0.15\tau_A$ and $t = 3\tau_A$ (dotted, dashed, dotted–dashed, and solid lines, respectively). In stark contrast to the highly nonlinear behavior of collisionless waves shown here, an MHD perturbation at these parameters is almost perfectly linear. A detailed description of each phase of evolution is given in section 5.1.

after the small-scale firehose modes have grown at the antinodes and returned the anisotropy to the marginal level. During this phase, the presence of a spatially varying nonlinearity—i.e., the $\mathcal{O}(\beta^{-1/2})$ spatial variation in Δ (equation (38)) and the $(1 + \delta b^2)^{-1}$ nonlinearity arising from field strength variation (see equation (20))—is crucial to the dynamics, because the linear term $1 + \beta\Delta/2$ is small. Without these nonlinearities, there is no preferential location for the growth of firehose instabilities and the entire wave erupts in a sea of small-scale fluctuations. It is critical (but nontrivial) to account for such a nonlinearity in a reduced equation that describes interruption dynamics (see appendix A.4, equation (A.60) and figure A2, for such an equation).

Late-time evolution. As the firehose modes push the anisotropy back to its marginal level, the smallest-scale fluctuations decay rapidly (Melville *et al* 2016). Following a transient period during which the pressure anisotropy slowly oscillates around (and decays towards) $\Delta = -2/\beta$ (Schekochihin *et al* 2008), the system relaxes into a final state with regions of straight fields separated by sudden corners. Despite this state's bizarre appearance, the basic cause of the plasma's preference for such structures may be inferred from a rather simple physical argument (within the LF model). This argument follows from three important properties of the collisionless dynamics: (i) without particle scattering any decrease in the magnetic field will lead to a decrease in the pressure anisotropy towards more negative values[17]; (ii) the only way in which the magnetic-field strength can be constant in time is either for the anisotropy to be at the firehose limit or for the field lines to be straight (or both); (iii) the heat fluxes continue to remove the spatial variation in Δ during the slow transient phase following the initial interruption. Property (i) tells us that the magnetic field cannot continue decreasing, as it did in the Braginskii model, without creating small-scale firehose fluctuations everywhere in the plasma. Then, if we assume that the plasma has reached some quasi-steady state with nonzero B, the plasma cannot be everywhere at the firehose limit and also have $\partial_z B \neq 0$, because the heat fluxes continue to flatten Δp (see figure 6(b)). Thus, in the absence of oscillatory behavior due to the Lorentz force, properties (ii), (iii) together imply that that B is effectively also flattened by the heat fluxes, which in turn suggests $\delta b\,(z)$ must be piecewise constant if it is nonzero. The result is the zig-zag field lines, shown in figure 6(c). Note that the double-adiabatic model, which neglects the heat fluxes and so lacks property (iii), does not produce constant B fields (see figure A1); instead, fields with curvature may be tensionless by being everywhere at the firehose limit but with a spatially varying Δp, as in the case of Braginskii interruption (see section 4.1).

[17] See equation (18) for the evolution of the spatially averaged anisotropy. The asymptotic scalings discussed in A.3 show that the compressional term is small at high β, while the heat-flux term, $3\langle q_\perp \nabla\cdot\hat{\boldsymbol{b}}\rangle$, relies on magnetic curvature and so will decrease with δb. Thus, the only effect able to cancel the creation of anisotropy through $3p_0\langle B^{-1}\mathrm{d}B/\mathrm{d}t\rangle$ is the collisional damping $-3\nu_c\langle\Delta p\rangle$.

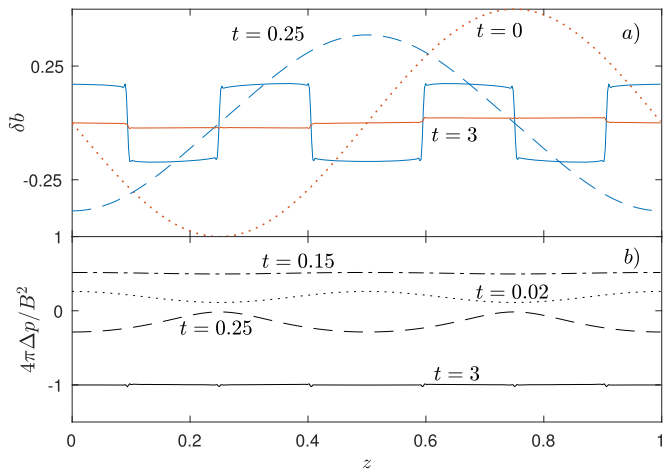

Figure 7. Evolution of a SA standing wave in the collisionless LF model at $\beta = 100$ (same parameters as figure 6), with an initial velocity perturbation above the interruption limit $u_\perp/v_A = -0.5\sin(k_z z)$ (i.e., initial conditions (22)). Panel (a) shows $u_\perp/v_A$ at $t = 0$ and $t = 3\tau_A$ (dotted and solid red lines respectively), and δb at $t = 0.25\tau_A$ (when δb is at its maximum) and $t = 3\tau_A$ (dashed and solid blue lines respectively). We limit $\Delta \leqslant 1/\beta$ to capture heuristically the anisotropy-limiting behavior of the mirror instability; this enables δb to reach amplitudes approaching that of the initial $u_\perp/v_A$. Panel (b) shows the anisotropy parameter $4\pi\Delta p/B^2$ at $t = 0.02$ (dotted line; this shows the early time increase in Δp), at $t = 0.15$ (dotted–dashed line; when the anisotropy is limited at $\Delta = 1/\beta$), $t = 0.25$ (dashed line), and $t = 3$ (solid line). Following the decrease in the magnetic field from the profile shown at $t = 0.25\tau_A$, the evolution is relatively similar to that shown in figure 6 (we do not illustrate intermediate times to avoid clutter).

It is worth noting that, within the LF model that we use (equations (3)–(7) with equations (10) and (11)), the firehose fluctuations grow fastest at the smallest scale accessible in the simulation, which is set by an artificial hyperdiffusion operator. In reality, this scale is set by the gyroradius, where the parallel firehose growth rate decreases due to finite-Larmor radius (FLR) effects (Davidson and Völk 1968, Schekochihin *et al* 2010). A more detailed study of FLR and other kinetic effects will be the subject of future work (see section 6), but it is worth noting that we see very similar macroscopic dynamics independently of the numerical resolution (for resolutions $N_z \gtrsim 128$). This suggests that the exact scale separation between the SA wave and the firehose fluctuations that erupt in the 'early nonlinear evolution' phase is not important for the late-time large-scale evolution (so long as the scale separation is sufficiently large).

5.1.1. Standing waves with an initial velocity perturbation

The discussion above concerned the evolution of a wave starting from a magnetic perturbation. For an initial perturbation in the velocity, the anisotropy initially grows in the positive direction, effectively increasing the restoring force of the wave. Starting from $\Delta p = 0$ but without a mechanism to limit Δp at positive values, this increase of Δp as B grows is exactly the same as its decrease after B has reached its maximum, so the system never reaches $\Delta < 0$. This results in nonlinear standing-wave oscillations with a frequency $\omega > \omega_A$ and $u_\perp/v_A > B_\perp/B_0$[18], which decay in time because of pressure-anisotropy damping arising from the (small) spatial variation in Δp (see section 5.2 for details). However, the mirror instability (which is excited when $\Delta > \beta^{-1}$) breaks this symmetry, allowing the magnetic field to grow in time while Δ is fixed at $\Delta \approx \beta^{-1}$. As the magnetic field starts decreasing again, the mirror modes that sustained $\Delta \approx \beta^{-1}$ presumably decay quickly (Melville *et al* 2016; see also our section 6), implying that the anisotropy starts decreasing from $\Delta = \beta^{-1}$ and can reach negative values. Thus, the limit on $u_\perp/v_A$ will be similar to equation (40), with perhaps a larger numerical prefactor to account for the fact that the magnetic-field decrease starts from a positive pressure anisotropy. The process is illustrated in figure 7, in which we artificially limit

[18] Using equation (38) and assuming that $\partial_z\Delta = 0$, it is straightforward to derive an equation for the amplitude δb of a (sinusoidal) perturbation. After normalizing time by ω_A, this is

$$\partial_t^2\delta b(t) = -\delta b - \frac{3\beta}{8}\delta b^3, \tag{41}$$

with the initial condition $\delta b(0) = 0$, $\partial_t\delta b(0) = \delta u_0 = u_\perp(0)/v_A$. Equation (41) is an undamped Duffing equation, and for $\beta^{1/2}\delta u_0 \gg 1$ has the approximate solution

$$\delta b(t) \approx \left(\frac{16\delta u_0^2}{3\beta}\right)^{1/4} \mathrm{sn}\left[\left(\frac{3\beta\delta u_0^2}{16}\right)^{1/4} t, -1\right], \tag{42}$$

where sn denotes the Jacobi elliptic function. These solutions oscillate in time, with a maximum amplitude $\delta b_{\max} < \delta u_0$ and a frequency larger then ω_A (this is 1 in these time units). There is no damping of the wave, which arises from the neglected spatial variation in the pressure anisotropy.

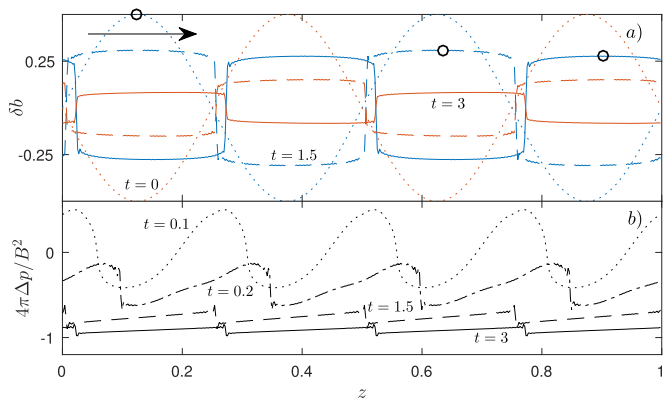

Figure 8. Evolution of a shear-Alfvén traveling wave in the collisionless LF model at $\beta = 100$, with an initial amplitude above the interruption limit, $\delta b = -u_\perp/v_A = 0.5\sin(4\pi z)$. Panel (a) illustrates δb (blue) and $u_\perp/v_A$ (red) at $t = 0$ (dotted lines), $t = 1.5\tau_A$ (dashed lines), and $t = 3\tau_A$ (solid lines). Panel (b) illustrates the anisotropy, $4\pi\Delta p/B^2 = \Delta\beta/2$, associated with the evolution shown in (a), at $t = 0.1\tau_A$, $t = 0.2\tau_A$, $t = 1.5\tau_A$ and $t = 3\tau_A$ (dotted, dashed, dotted–dashed, and solid lines respectively). Note the slow decrease in mean anisotropy, which forces the magnetic field to dominate over the velocity and slows the wave. At later times (not shown due to clutter in the figure), the velocity continues to damp and the wave eventually comes to a standstill, with a similar final state to that of the standing waves shown in figures 6 and 7.

the anisotropy to $\Delta \leqslant \beta^{-1}$. Although $\Delta = \beta^{-1} > 0$ when the magnetic perturbation reaches its maximum (at $t \approx \tau_A/4$)[19], the resulting final state is similar to figure 6.

5.2. Traveling waves

An unperturbed traveling wave, with $\delta b \propto \cos(k_z z - \omega_A t)$, does not change $\langle B(t)\rangle$, because the wave simply shifts across the domain as it evolves (unlike a standing wave, with $\delta b \propto \cos(k_z z)\cos(\omega_A t)$). The arguments developed for standing waves in the previous section thus no longer apply, since the $\mathcal{O}(1)$ part of equation (38) is zero. However, even though $\langle B^{-1}\mathrm{d}B/\mathrm{d}t\rangle = 0$, $B^{-1}\mathrm{d}B/\mathrm{d}t$ itself is large. The resulting Δp—which is reduced by a factor $\sim\beta^{1/2}$ by the heat fluxes —is correlated in space with $B^{-1}\mathrm{d}B/\mathrm{d}t$ and thus damps wave energy into thermal energy at the rate (17), viz.,

$$\partial_t E_{\mathrm{wave}} = -\int \mathrm{d}\boldsymbol{x}\ \Delta p \frac{1}{B}\frac{\mathrm{d}B}{\mathrm{d}t}. \tag{43}$$

As the wave is damped, the resulting decrease in $\langle B(t)\rangle$ causes $\langle\Delta p\rangle$ to decrease also (according to equation (38)), slowing down the wave and eventually causing it to stop (interrupt) if the initial B is sufficiently large for $\langle\Delta\rangle$ to reach $-2/\beta$. The key point here is that during the process of wave decay, there is no mechanism to isotropize the $k_\parallel = 0$ component of the pressure, implying any decrease in B must also be accompanied by a decrease in Δp.

The process described above is illustrated in figure 8, which shows the solution of the LF equations at the same parameters as the standing-wave examples (figures 6 and 7). At early times, the pressure anisotropy (the dotted line in figure 8(b)) is a strong function of space[20] with $\langle\Delta p\rangle = 0$. This spatially periodic Δp then damps the wave, as well as causing significant nonlinear modifications to its shape (which becomes more angular, reducing $B^{-1}\mathrm{d}B/\mathrm{d}t$). As is also clear in figure 8(b), this damping drives the mean anisotropy to negative values. This effectively reduces the Alfvén speed to $\tilde{v}_A = v_A(1 + \beta\Delta/2)^{1/2}$, which causes the velocity to decay faster than the magnetic field (compare red and blue dashed and solid lines in figure 8(a)) because $\delta u_\perp/\tilde{v}_A = \delta B_\perp/B_0$ in a traveling wave and $\tilde{v}_A < v_A$. Although not shown in figure 8 to avoid clutter, the velocity continues to be damped faster than the field as Δ approaches $-2/\beta$, and we are left with an angular, magnetically dominated final state that is similar to the final state of the standing-wave evolution (figures 6 and 7).

5.2.1. Landau damping versus pressure-anisotropy damping

It is worth noting that the Landau damping rate of a linearly polarized SA wave due to the (nonlinear) spatiotemporal variation of the magnetic pressure is similar to the pressure-anisotropy damping, causing the wave to be damped at the rate $\gamma/\omega_A \sim \beta^{1/2}\delta b^2$ (Hollweg 1971b, Lee and Völk 1973; see discussion around

[19] The magnetic perturbation $\delta B_\perp/B_0$ that results from $u_\perp$ is also slightly smaller than the initial $u_\perp/v_A$ because of the larger restoring force when $\Delta > 0$.

[20] Note that the smoothing effect of the heat fluxes is very strong here. Without it, the early time anisotropy shown in figure 8(b) would be much larger. In the example shown in figure 8, the early time anisotropy is below the mirror and firehose thresholds ($|\Delta| \lesssim \beta^{-1}$); however, for even larger amplitudes, $\delta b \gtrsim \beta^{-1/4}$ the wave can cause $|\Delta| \gtrsim \beta^{-1}$ in the regions where the field is changing fastest (this estimate results from $|\Delta|_{\mathrm{max}} \sim \beta^{-1/2}\delta b^2$; see equation (A.43)), causing even stronger nonlinear modifications to the wave.

equation (48) below)[21]. This effect is different from pressure-anisotropy damping (but is still captured by the LF model), and can in fact also be included in models of wave propagation that do not include a pressure anisotropy (see, e.g., Medvedev and Diamond 1996, Medvedev *et al* 1997). The difference between the pressure-anisotropy damping and Landau damping can be explained as follows. Heat fluxes are necessary for the Landau damping of a nonlinear wave, because they directly damp out the pressure perturbation that arises from the magnetic-field-strength variation. In contrast, heat fluxes act to *reduce* the pressure-anisotropy damping, by smoothing the spatial variation in the pressure anisotropy (in other words, the Landau damping damps the pressure-anisotropy damping!). In the discussion below, our estimate of the damping rate is heuristic, so there is no need to work out these two effects separately; however, Landau-damping effects are included in the calculation of the wave-decay rate given in appendix A.3.

5.2.2. Semi-quantitative description of traveling wave evolution

We now analyze the traveling-wave decay process in more detail, deriving a simple ordinary differential equation (ODE) to describe the process of decay and interruption. We assume that the wave remains sinusoidal throughout its evolution, which, although far from quantitatively justified (see figure 8), allows one to construct an ODE that describes qualitatively how the nonlinear damping and mean anisotropy affect the magnetic-energy decay. This is then used to derive the decay rates of kinetic and magnetic energy,

$$\gamma_U \equiv \frac{1}{2E_K}\frac{\mathrm{d}E_K}{\mathrm{d}t}, \qquad \gamma_B \equiv \frac{1}{2E_M}\frac{\mathrm{d}E_M}{\mathrm{d}t}, \tag{44}$$

as a function of β and δb. These turn out to match reasonably well the numerical LF solutions.

Our first step is to work out the decay rate of the wave due to pressure-anisotropy and Landau damping. Because this relies on the LF prescription for the heat fluxes, which in turn depends on the compressible response of the plasma (because $q_{\perp,\parallel} \sim \partial_z(p_{\perp,\parallel}/\rho)$), a formal calculation of this damping is somewhat involved and is worked out in detail in appendix A.3 (see equations (A.44) and (A.45)). Here we give a heuristic derivation so as to present a relatively simple description of the important physics. Defining $\Delta = \bar{\Delta} + \Delta_k$, where $\bar{\Delta} = \langle\Delta\rangle$ and Δ_k is the spatially varying part of Δ that arises from the perturbation $\delta b(z)$ with wavenumber k, the important terms in the equation for Δ_k are (see equation (13))

$$\partial_t \Delta_k \approx 3\frac{1}{B}\frac{\mathrm{d}B}{\mathrm{d}t} - a_1\sqrt{\frac{p_0}{\rho}}\,|k_\parallel|\Delta_k, \tag{45}$$

where a_1 is an $\mathcal{O}(1)$ dimensionless coefficient, which depends on the details of a closure for the heat fluxes. We assume a monochromatic traveling wave of amplitude δb, giving

$$2\frac{1}{B}\frac{\mathrm{d}B}{\mathrm{d}t} \approx \frac{3}{2}\delta b^2\omega_A(\bar{\Delta})\sin[2kz - 2\omega_A(\bar{\Delta})t], \tag{46}$$

where $\omega_A(\bar{\Delta}) = v_A k(1 + \beta\bar{\Delta}/2)^{1/2}$ accounts for the slowing down of the wave as $\bar{\Delta}$ becomes negative. Since $k_\parallel c_s \sim \beta^{1/2}\omega_A$, we may neglect the time derivative[22] on the left-hand side of equation (45), leading to

$$\Delta_k \sim \beta^{-1/2}\delta b^2\left(1 + \beta\frac{\bar{\Delta}}{2}\right)^{1/2}\sin[2kz - 2\omega_A(\bar{\Delta})t]. \tag{47}$$

Evaluating the integral (43) one finds,

$$\partial_t E_{\mathrm{wave}} \approx a_2\beta^{1/2}B_0^2\delta b^4\omega_A\left(1 + \frac{\beta\bar{\Delta}}{2}\right), \tag{48}$$

where $a_2 \approx \sqrt{\pi}/8$ is calculated in appendix A.3 (equation (A.45)). This expression is similar in form to the damping rate (36) in the Braginskii regime, but is reduced by $\beta^{1/2}$ due to the smoothing effect of the heat fluxes.

Armed with the energy damping rate (48), we now formulate an equation for the slow (compared to the sound propagation, $\omega \sim k_\parallel c_s$) dynamics of the wave amplitude δb. For a sinusoidal wave, the magnetic energy is $E_M = \delta b^2 B_0^2/16\pi$, while the assumption that the wave remains traveling rather than standing (i.e., that it does not generate a global oscillation in time) gives $E_K = (1 + \beta\bar{\Delta}/2)E_M$. Noting also that $\bar{\Delta} = 3/4(\delta b^2 - \delta b_0^2)$, equation (48) becomes

[21] This estimate neglects particle trapping effects, which reduce the damping rate (Kulsrud 1978, Cesarsky and Kulsrud 1981, Völk and Cesarsky 1982); i.e., the damping-rate estimate used here results from an application of linear Landau damping to a nonlinear wave. Such particle trapping effects are not included in our model, because our LF closure uses linear Landau-damping rates to calculate the heat fluxes.

[22] The heat fluxes suppress spatial variation in Δ_k on the time scale $\tau_{\mathrm{damp}} \sim (k_\parallel c_s)^{-1} \ll |\nabla \boldsymbol{u}|^{-1} \sim \omega_A$, so $\partial_t\Delta \sim \omega_A\Delta$ is small compared to $|k_\parallel|c_s\Delta$.

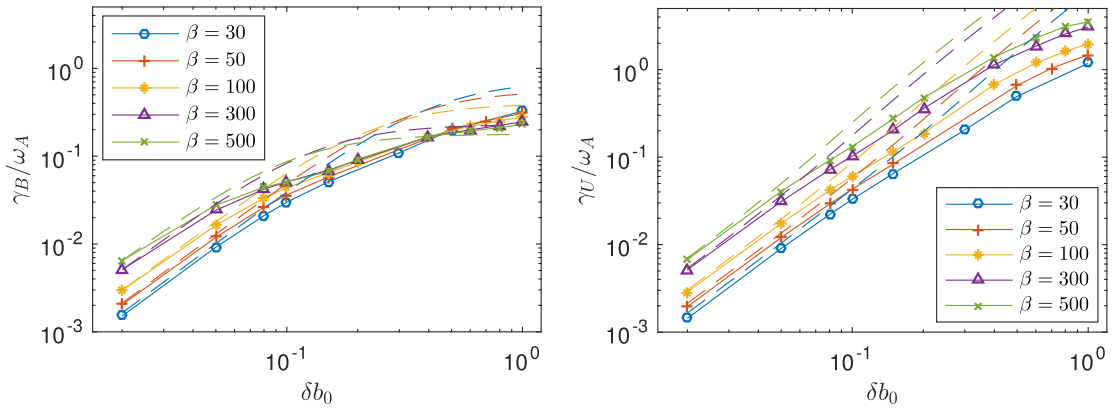

Figure 9. Initial traveling-wave decay rates. The solid lines and symbols show decay rates from the first $2\tau_A$ of the wave's evolution, measured in numerical simulation of the LF equations with initial conditions $\delta b(z, 0) = -\delta b_0 \cos(k_z z)$, $u_x(z, 0)/v_A = \delta b_0 \cos(k_z z)$. The dashed lines show the theoretical predictions (52). The left panel shows the magnetic-field decay rate γ_B/ω_A while the right panel shows the velocity decay rate γ_U/ω_A (see equation (44)). Although the agreement is not perfect, the theoretical predictions do capture the qualitative trends rather well, considering the crudeness of the sinusoidal approximation used.

$$\partial_t\left[2\delta b^2 + \frac{3\beta}{8}\delta b^2(\delta b^2 - \delta b_0^2)\right] = -a_2\beta^{1/2}\omega_A\delta b^4\left[1 + \frac{3\beta}{8}(\delta b^2 - \delta b_0^2)\right]. \tag{49}$$

This can be reformulated in the variables $\zeta = \beta\delta b^2$ and $\bar{t} = \omega_A\beta^{1/2}t$ as

$$\partial_t\left[2\zeta + \frac{3}{8}\zeta(\zeta - \zeta_0)\right] = -a_2\zeta^2\left[1 + \frac{3}{8}(\zeta^2 - \zeta_0^2)\right], \tag{50}$$

which has the benefit of being controlled by just one parameter $\zeta_0 = \beta\delta b_0^2$.

A full analytic solution to equation (50) is intractable, but numerical solutions (not shown) match our expectations based on the qualitative discussion in section 5.2 above. Specifically, above the interruption limit ($\zeta_0 \gtrsim 1$), $u_\perp$ decays much faster than $\delta B_\perp$ in time and $\delta B_\perp$ asymptotes to a constant nonzero value at late times, whereas below the interruption limit ($\zeta_0 \lesssim 1$) the nonlinear damping more equally affects $u_\perp$ and $\delta B_\perp$ and there is simply a slow decay of both.

We now use equation (50) to derive the initial decay rates of $u_\perp$ and $\delta B_\perp$. This is most easily done by linearizing equation (50) about $\zeta = \zeta_0$, viz., letting $\zeta = \zeta_0(1 + \delta\zeta)$ and expanding in $\delta\zeta$. This gives

$$\delta\zeta = -\frac{a_2\zeta_0}{2 + 3\zeta_0/8}\bar{t}. \tag{51}$$

Rewriting ζ in terms of δb, substituting $\bar{t} = \omega_A\beta^{1/2}t$, then calculating the decay rates (44) gives

$$\gamma_B \approx \frac{1}{2}a_2\omega_A\beta^{1/2}\delta b_0^2\frac{8}{16 + 3\beta\delta b_0^2}, \qquad \gamma_U \approx \frac{1}{2}a_2\omega_A\beta^{1/2}\delta b_0^2\frac{8 + 3\beta\delta b_0^2}{16 + 3\beta\delta b_0^2}. \tag{52}$$

Although these expressions appear rather complicated, they agree nicely with our intuitive picture described earlier. In particular, the decay transitions from a regime where $\gamma_B \approx \gamma_U$ below the interruption limit $\beta\delta b_0^2 \ll 1$, to one where $\gamma_B \ll \gamma_U$ (with γ_B independent of δb_0) when $\beta\delta b_0^2 \gg 1$.

A comparison of the damping rates (52) to the full LF traveling-wave solutions is presented in figure 9, where we show the decay rates measured numerically for solutions starting with a sinusoidal traveling wave. We find good agreement with the damping rates at low δb_0, when they are small, and qualitative agreement with the trends predicted by equation (52) at larger δb_0. Note in particular that the decay rate of $\delta B_\perp$ changes from increasing to decreasing with β at high δb_0, whereas the decay rate of $u_\perp$ does not. The quantitative agreement at high $\beta\delta b_0^2$ is lacking, and there are clear reasons for this discrepancy. First, there is our assumption that the wave remains sinusoidal, which is patently not true when $\beta\delta b_0^2 > 1$ (see figure 8). The strong nonlinear shape modifications that do occur early in the evolution presumably involve some exchange of energy between $u_\perp$ and $B_\perp$, in ways that are not included in our model. Secondly, the measurement of a decay rate is ambiguous for the strongly nonlinear $\beta\delta b_0^2 > 1$ solutions. For simplicity, we have fit the amplitude evolution from $t = 0$ to $t = 2\tau_A$ to a decaying exponential function, but the decay rate can vary significantly over this range at $\beta\delta b_0^2 > 1$. We have explored a variety of methods for determining this initial decay rate of the wave, and although the quantitative results vary with method, the general properties and qualitative agreement with the predictions (52) are robust.

6. Fully kinetic and multi-dimensional effects

Throughout the preceding sections, we have primarily focused on physical effects contained within the simplest 1D LF equations (3)–(11). Importantly, the mirror and oblique firehose instabilities are not included in this model[23], because these grow at $k_\perp \sim k_\parallel$ on the Larmor scale and thus require 2D or 3D kinetic simulations to be resolved correctly. In this section, we discuss—based on previous fully kinetic theory and simulations—some possible effects of these microinstabilities on the global wave evolution, focusing on which aspects of the simple 1D picture described above are robust, and which may be modified by the inclusion of this physics. We also discuss other kinetic effects that could modify our results, including FLR effects (these were neglected by assuming $k_\parallel \rho_i \ll 1$), particle trapping (this is not contained with the LF prescription for the heat fluxes), heat-flux limits from the gyrothermal instability, and other scattering effects. Of course, this discussion is in no way intended to be a replacement for future fully kinetic theory and simulations in two or three dimensions; rather, its purpose is to motivate the design of such studies and provide some guidance for interpreting their results.

6.1. Mirror instability

In our discussion of standing waves (section 5.1), the mirror instability was invoked to justify a limit on positive pressure anisotropies when starting from a velocity perturbation. This in turn allowed the magnetic field to grow, reach its maximum, and then be interrupted in a similar way to an initial purely magnetic perturbation. Without this limiting effect, an initial velocity perturbation will create an oscillating wave (albeit not a linear SA wave because the restoring force is enhanced by the positive anisotropy; see equations (41) and (42)). Thus, although the mirror instability is not crucial for the interruption effect itself, its presence does imply that the effect cannot be significantly modified based on the initial conditions. We now discuss in more detail why it is reasonable to assume that the mirror instability should have this effect.

A variety of recent kinetic results (Schekochihin *et al* 2008, Kunz *et al* 2014, Hellinger and Trávníček 2015, Rincon *et al* 2015, Melville *et al* 2016, Riquelme *et al* 2016) show that mirror fluctuations, which are unstable when $\Delta \gtrsim \beta^{-1}$ and cause perturbations in the field strength δB, limit Δp by trapping particles. Namely, as the macroscopic field grows and attempts to raise the pressure anisotropy, a larger and larger fraction of particles becomes trapped in the magnetic wells and 'sees' a lower field. Thus, even though the volume-averaged field continues to increase, Δ is maintained at the marginal level β^{-1} because a larger proportion of particles is trapped in the ever-deepening mirror wells. During this phase, the magnetic mirrors grow in time as $|\delta B/B| \sim (|\nabla \boldsymbol{u}|t)^{2/3}$ and there is very little particle scattering because their parallel scale is significantly larger than the Larmor radius. Further, since the mirrors only saturate and start scattering particles when $|\delta B/B| \sim 1$, they should never saturate and cause significant particle scattering for any SA wave initial conditions with $u_x(0)/v_A \sim |\nabla \boldsymbol{u}|\tau_A < 1$. The numerical experiments and arguments of Melville *et al* (2016) are particularly relevant to what happens as the magnetic field reaches its maximum and starts to decrease. For $\beta \ll \Omega_i/|\nabla \boldsymbol{u}|$—the 'moderate-$\beta$', or large-scale-separation, regime most relevant to our results—the mirrors should freely decay on time scales much shorter than τ_A (the decay time is $\sim \beta/\Omega_i$), releasing their trapped particles and allowing the anisotropy to decrease towards the firehose limit. Although less well understood, it seems that in the opposite limit, $\beta \gg \Omega_i/|\nabla \boldsymbol{u}|$, the firehose limit is *also* quickly reached (see Melville *et al* 2016, section 3.2), probably because the smaller-scale firehose fluctuations are able to grow on top of the larger-scale decaying mirrors. Overall, it is thus reasonable to surmise that the mirror instability will effectively act as a passive limiter, ensuring $\Delta \lesssim \beta^{-1}$ but not strongly affecting large-scale wave dynamics[24].

It is worth reiterating a fundamental difference between the mirror and firehose limits for SA waves. At the firehose limit, the anisotropic stress nullifies the wave restoring force (i.e., the magnetic tension). In contrast, a plasma at the mirror limit merely feels a modestly stronger (factor 3/2) restoring force. This difference explains why the firehose limit is of much greater importance than the mirror limit for SA wave dynamics.

6.2. Oblique firehose instability

The oblique firehose instability is not included in our model, both because it operates at $k_\perp \sim k_\parallel$ and because kinetic theory is required for its correct description (Hunana and Zank 2017). Linearly, oblique firehose fluctuations grow faster than parallel firehose fluctuations (growth rate $\gamma_{\max} \sim |\Delta + 2/\beta|^{1/2}\Omega_i$, as opposed to

[23] The linear mirror instability can be captured relatively accurately by the LF model that we use (if the equations are solved in 2 or 3 dimensions; see section 8 of Snyder *et al* 1997). However, given the importance of trapped particles in the nonlinear mirror evolution (Schekochihin *et al* 2008, Kunz *et al* 2014, Rincon *et al* 2015, Melville *et al* 2016), it seems quite unlikely that a LF model could correctly reproduce the pressure-anisotropy-limiting behavior of the mirror instability, although we know of no relevant study that tests this (see section 6.1 for discussion).

[24] As the scale separation $(k_\parallel \rho_i)^{-1}$ is reduced, the mirrors will presumably become less effective (see Kunz *et al* 2014), allowing Δ to overshoot β^{-1} before acting to limit the anisotropy. Thus very large domains (compared to ρ_i) are likely essential to see these effects in fully kinetic simulations.

$\gamma_{\max} \sim |\Delta + 2/\beta|\Omega_i$ for the parallel firehose) because of their smaller scale (Yoon *et al* 1993, Hellinger and Matsumoto 2000, Hellinger and Trávníček 2008, Rosin *et al* 2011), and are also seen clearly at larger amplitudes in nonlinear regimes at high β (Kunz *et al* 2014, Melville *et al* 2016). Further, unlike mirror fluctuations, the firehose fluctuations will saturate and start scattering particles after $t_{\rm sat} \sim \beta^{1/2}(|\nabla \boldsymbol{u}|\Omega)^{-1/2}$ (for $\beta \ll \Omega/|\nabla \boldsymbol{u}|$; see Kunz *et al* 2014, Melville *et al* 2016); i.e., after a very short time set by the microphysics.

The most obvious question that arises is then whether the differences between oblique and parallel firehose dynamics will cause significant differences in the nonlinear interruption of SA waves, compared to 1D models where only the parallel firehose exists. This remains unclear, and understanding such issues will require fully kinetic simulations in 2D or 3D with $k_z \rho_i \ll 1$. Either scenario—that the oblique firehose does or does not significantly modify the SA wave dynamics—can be plausible. On the one hand, the oblique firehose may behave similarly to the parallel firehose in 1D simulations: be strongly excited during the early phases of wave interruption, but then die away at later times because the pressure anisotropy is pushed back above the firehose limit. On the other hand, the enhanced particle scattering in kinetic oblique firehose fluctuations could possibly continue until the magnetic field decays completely, potentially leading to collisionless SA dynamics that more closely resemble a Alfvén wave in the Braginskii regime[25] (see section 4). However, in either case, the presence of oblique firehose fluctuations cannot circumvent the interruption limit itself—they have approximately same instability threshold as the parallel firehose (Hellinger and Matsumoto 2000, Klein and Howes 2015), so will become important only after the wave restoring force has already disappeared.

6.3. Other kinetic effects

Here we outline several other possible kinetic effects. Unlike the mirror and oblique firehose modes discussed above, most of these effects could be studied using 1D, but fully kinetic simulations.

FLR effects. Although various FLR effects can be included in LF models (Goswami *et al* 2005, Ramos 2005, Passot *et al* 2012, Sulem and Passot 2015), the simple closure that we used here does not include these corrections. Assuming large scale separation, the most obvious effect from such corrections in 1D is the regularization of the small scales for the parallel firehose, which has its peak growth rate at $k_\parallel \rho_i \sim |\Delta + 2/\beta|^{1/2}$ (Davidson and Völk 1968, Schekochihin *et al* 2010). Since we found that collisionless wave-interruption dynamics did not depend significantly on numerical resolution (which effectively sets the fastest-growing firehose mode in our fluid model), it seems unlikely that the direct effect of this regularization will be particularly important to the large-scale interruption (but note that the requirement $k_\parallel \rho_i \ll 1$ could be quite severe, because $|\Delta + 2/\beta|^{1/2} \ll 1$ and we need significant separation between the firehose modes and the wave). There could, however, be other effects that are of some significance. For example, FLR effects enable a new instability—the 'gyrothermal instability' (Schekochihin *et al* 2010)—which may act to limit the heat fluxes before the SA wave hits the interruption limit, in a similar way to how the firehose instability limits the pressure anisotropy (Rosin *et al* 2011). Through the gyro-viscous terms (the off-diagonal elements of the pressure tensor) and the Hall effect[26], FLR effects can also act to circularly polarize the wave, creating a B_y perturbation from a spatially varying B_x. However, this is presumably only directly important for the macroscopic wave when the scale separation is modest, or in regions with large gradients that form during nonlinear evolution. Experimenting with simple extensions to the LF model that include gyro-viscous effects and/or the Hall effect (not shown here), we have seen that these terms cause only minor changes to the SA wave evolution, so long as $k_\parallel \rho_i$ is sufficiently small. Note that the Spitzer resistivity, which arises from electron ion collisions and contributes a term of the form $\eta \nabla^2 \boldsymbol{B}$ to the induction equation (5), is negligibly small for the weakly collisional plasmas we consider here, and cannot provide significant smoothing of the magnetic field above the ion gyroscale[27].

Particle trapping. Since LF closures prescribe the heat fluxes based on linear Landau-damping rates, effects of particle trapping are not included in these closures and may provide an order-unity correction to the heat fluxes. In particular, trapping can be important whenever the bounce frequency ω_b of particles approaches the frequency of large-scale motions (this is $\sim\omega_A$ for a SA wave). Given that particles with velocity $\boldsymbol{v}$ and parallel velocity $v_\parallel$ are trapped if $\xi = v_\parallel/|\boldsymbol{v}| < \xi_{\rm tr} \sim |\delta B/B_0|^{1/2}$, while $|\delta B/B_0|^{1/2} \sim |\delta B_\perp/B_0|$, a simple estimate for the

[25] In support of this idea, including an artificial hard-wall firehose limit at $\Delta = -2/\beta$ in a standing-wave LF simulation leads to the wave being strongly nonlinearly modified and then rather quickly decaying to oscillate with an amplitude below the interruption limit. The effect on a traveling wave is less severe, because the wave remains above the firehose limit for much of its decay.

[26] The Hall effect becomes important when $k_\parallel d_i \sim 1$, where $d_i \sim \sqrt{\beta}\rho_i$ is the ion skin depth.

[27] For example, computing the diffusion time at the gyroscale due to resistivity, $(\eta/\rho_i^2)^{-1}$, for fiducial solar wind parameters (Montgomery 1983, Bruno and Carbone 2013), one obtains

$$\left(\frac{\eta}{\rho_i^2}\right)^{-1} \approx 2 \times 10^6 \left(\frac{B}{6\ \mathrm{nT}}\right)^{-1} \left(\frac{T}{10^5 \mathrm{K}}\right)^{5/2} \Omega_i^{-1}. \tag{53}$$

This very long timescale implies that the effect of resistivity is negligible for magnetic field variations at (or above) the ion gyroscale, and other FLR effects (e.g., the Hall effect) will cause more significant modifications to the magnetic field evolution.

bounce frequency is $\omega_b \sim \beta^{1/2}\delta b\ \omega_A$. Thus, trapping can be important if $\delta b > \beta^{-1/2}$; i.e., for a wave above the interruption limit. Trapping has the effect of reducing the Landau damping rate of nonlinear SA traveling waves (O'Neil 1965, Kulsrud 1978) and presumably also modifies pressure-anisotropy damping. These effects will be considered in more detail in future work.

Other scattering effects. If the 'corners' that develop in the magnetic-field lines (e.g., figures 6–8) are on the Larmor scale, unresolved by our LF closure, these may scatter particles. This would provide an interesting case where a plasma could set its own mean free path $\lambda_{\rm mfp} \sim k_\parallel^{-1}$ based on the large-scale driving. If real, this effect would most significantly modify traveling-wave dynamics, because the square structures that develop *before* interruption (see figure 8) could possibly cause sufficient scattering to damp the global pressure anisotropy fast enough so that the wave decayed before reaching the interruption limit.

Overall, we would like to stress that although the details of wave interruption may be modified by the addition of other kinetic physics, our basic result—that weakly collisional SA waves cannot exist in their linear form above the limits (30) and (40)—is robust. The dominance of magnetic energy over kinetic energy is also a generic consequence of interruption, because in the approach to the firehose limit, the equipartition of energy in an Alfvén wave is modified by the decrease of magnetic tension. We find generic agreement on these points between the LF, Braginskii, and double-adiabatic models ($q_\perp = q_\parallel = 0$; see appendix A.2). Many of our key results are thus quite insensitive to the form of the heat fluxes or particle scattering, relying purely on the physics of pressure-anisotropy generation in a changing magnetic field.

7. Conclusion

In this paper, we have explored the nonlinear 'interruption' and damping of linearly polarized SA waves in weakly collisional plasmas. These effects, which arise due to the pressure anisotropy that is generated in the changing magnetic field of the wave, lead to a limit on the amplitude of propagating/oscillating SA waves in the collisionless regime:

$$\frac{\delta B_\perp}{B_0} \lesssim \beta^{-1/2}. \tag{54}$$

In the weakly collisional Braginskii limit, which applies when $\nu_c \gg \omega_A$, propagating/oscillating SA waves are also limited in amplitude, to

$$\frac{\delta B_\perp}{B_0} \lesssim \sqrt{\frac{\nu_c}{\omega_A}}\,\beta^{-1/2}. \tag{55}$$

We summarize our main findings as follows.

- Above the limit (54), collisionless SA waves are 'interrupted' when their self-generated pressure anisotropy reaches the firehose boundary $\Delta p \approx -B^2/4\pi$. At this boundary, the wave's restoring force (the Lorentz force) is cancelled by the anisotropy, and the magnetic energy dominates the kinetic energy because the effective Alfvén speed goes to zero.
- Due to the correlation between Δp and $B^{-1}{\rm d}B/{\rm d}t$, there is a net transfer of wave energy to thermal energy of the plasma through 'pressure-anisotropy heating' at the rate $\int {\rm d}\boldsymbol{x}\ \Delta p B^{-1}{\rm d}B/{\rm d}t$. This results in a nonlinear damping of the wave, even below the limit (54).
- Heat fluxes are always important in the high-β collisionless limit (because the thermal velocity is larger than v_A) and act to smooth the spatial dependence of the pressure anisotropy.
- In the collisionless limit, standing and traveling SA waves behave in qualitatively different ways because the spatial average of B decreases during a standing wave's evolution, whereas it does not for a traveling wave. Thus, while a standing wave above the limit (54) is interrupted within half a wave period, a traveling wave is first nonlinearly damped (at the rate $\sim\omega_A(\delta B_\perp/B_0)^2\beta^{1/2}$), leading to a decreasing B and eventual interruption of the wave.
- The kinetic energy in a collisionless traveling wave is damped significantly faster than the magnetic energy for amplitudes approaching (or exceeding) the limit (54), and the magnetic energy can be a large fraction of its initial value when the wave interrupts. This occurs because, as the wave decays, the global decrease in Δp reduces v_A, which changes the ratio of $u_\perp$ and $\delta B_\perp$ (this also slows down the wave; see figure 8).
- Barring additional kinetic and higher-dimensional effects not contained within our LF model (see section 6), the outcome of wave interruption is the creation of a magnetically dominated state of nearly perfect zig-zag magnetic field lines (see figures 6 and 8)—i.e., a quasi-periodic pattern with spatially constant magnetic field

strength. The emergence of this state may be understood by noting that it is the only state that has both zero magnetic tension and a spatially smooth pressure anisotropy along the field lines (because spatial variation in Δp is damped by the heat fluxes).

- Wave interruption in the Braginskii limit involves a slow decay of the wave over the timescale $t_{\rm decay} \sim \beta/\nu_c(\delta B_\perp/B_0)^2$, which occurs because a slowly changing B is necessary to maintain the anisotropy at the firehose limit. The characteristic field-line structures (figure 2) differ from collisionless waves because the magnetic tension is zero if the anisotropy is at the firehose limit, even if there is spatial variation in B^2.
- The pressure-anisotropy damping of SA waves is large in a Braginskii plasma because the spatially varying part of Δp is comparable to its mean. For waves below the limit (55), this leads to the wave energy being damped at the rate[28] $\sim\omega_A^2/\nu_c(\delta B_\perp/B_0)^2\beta$.
- The amplitude limits do not apply to circularly polarized SA waves because for these, $\mathrm{d}B/\mathrm{d}t = 0$.

7.1. Implications and applications

Given the ubiquity of SA waves in plasma physics (see section 1), the stringent limits on their amplitude derived here may have interesting implications in a variety of hot, low-density (and therefore, weakly collisional) astrophysical plasmas. Although a detailed study of all these implications is beyond the scope of this work, it is worth commenting on magnetized turbulence in particular, given its importance in many subdisciplines of astrophysics. The salient point is that in well-accepted phenomenologies of strong magnetized turbulence (in particular, Goldreich and Sridhar 1995, and extensions, e.g., Boldyrev 2006), the physics of SA waves is critical at all scales in the turbulence[29]. A strong modification to SA wave physics would thus be expected to significantly modify the turbulent cascade. One might expect such modifications to be even stronger in the weak turbulence regime (Ng and Bhattacharjee 1996, Galtier *et al* 2000, Schekochihin *et al* 2012) given the relative weakness of nonlinear interactions in comparison to the SA wave physics in such turbulence.

More explicitly, turbulence in a weakly collisional high-β plasma may depend on the amplitude of its forcing. Since velocities are strongly damped when a wave is interrupted, it may behave as a fluid with Reynolds number $\lesssim 1$ when $u_\perp/v_A \gtrsim \beta^{-1/2}$ (and $u_\perp \lesssim v_A$; otherwise the waves are of such large amplitude that the turbulence would likely be in a dynamo regime, which is not particularly well understood even in MHD, but is less obviously Alfvénic). In contrast, for perturbations of amplitude $u_\perp/v_A \lesssim \beta^{-1/2}$, the linear SA wave physics is mostly unaffected, and a standard Alfvénic cascade should develop. Since pressure-anisotropy heating is able to dissipate large-scale wave energy when wave interruption is important, a turbulent cascade may not be necessary for the plasma to absorb the energy injected by a continuous mechanical forcing (Kunz *et al* 2010). While further study is necessary to understand this physics better, it is at least clear that the immediate interruption of SA fluctuations with amplitudes exceeding $\delta B_\perp/B_0 \sim \beta^{-1/2}$ should significantly limit the application of MHD-based turbulence phenomenologies to high-β weakly collisional plasmas.

7.2. Future work

A first priority for future studies of wave interruption is the inclusion of the kinetic effects discussed in section 6. Unfortunately (from a computational standpoint), due to the 2D kinetic nature of the mirror and oblique firehose instabilities, significant progress in this endeavor requires kinetic simulations in two spatial dimensions and three velocity-space dimensions. Since firehose instabilities grow on scales somewhat above the gyroscale (see section 6), if one hopes to study the asymptotic regime $k_\parallel\rho_i \ll 1$, the required scale separation between the gyroscale and the SA wave is likely quite large. We thus expect a detailed kinetic study to be rather computationally expensive, although certainly feasible. That said, there will also be a variety of interesting insights to be gained from purely 1D kinetics: for example, the role of the gyrothermal instability, particle scattering off magnetic discontinuities, and particle-trapping effects. Further, the behavior of SA waves with limited scale separation between $k_\parallel^{-1}$ and ρ_i is also of interest physically, in particular for the solar wind, where observations easily probe turbulent fluctuations down to the ion gyroscales and below. A separate line of investigation for future work involves applications of the amplitude limit, in particular to turbulence, as discussed in section 7.1. This may be productively pursued using a 3D LF code (as in Sharma *et al* 2006, 2007) or using Braginskii MHD.

Overall, although many questions remain, both the limit $\delta B_\perp/B_0 \lesssim \beta^{-1/2}$ itself and the strong dominance of magnetic over kinetic energy are robust, appearing across a variety of models. Given the stringent nature of the

[28] This result is valid in the regime where the heat fluxes are unimportant; see appendix B.

[29] A more common way to say this is that the cascade is in *critical balance* (Goldreich and Sridhar 1995), which states that the linear (Alfvén) time is equal to the nonlinear turnover time at all scales in a strong MHD cascade.

amplitude limit and the interesting implications for high-β magnetized turbulence in weakly collisional plasmas, we anticipate a range of future applications to heliospheric, astrophysical, and possibly laboratory (Forest *et al* 2015, Gekelman *et al* 2016) plasmas.

Acknowledgments

It is a pleasure to thank S Cowley, M Kunz, S Bale, C H K Chen, S Balbus, L Sironi, F Rincon, and M Strumik for useful and enlightening discussions. JS is endebted to Merton College for supporting a stay in Oxford where some of this work was carried out. JS and AAS would also like to thank the Wolfgang Pauli Institute in Vienna for its hospitality during the 9th Plasma Kinetics Working Group Meeting. JS was funded in part by the Gordon and Betty Moore Foundation through Grant GBMF5076 to Lars Bildsten, Eliot Quataert and E Sterl Phinney. AAS was supported in part by grants from UK STFC and EPSRC. EQ was supported by Simons Investigator awards from the Simons Foundation and NSF grant AST 13-33612.

Appendix A. Asymptotic wave equations—collisionless limit

In this appendix, we derive a variety of wave equations to describe standing and traveling SA waves in collisionless regimes (the Braginskii regime is treated in appendix B). This is carried out by means of asymptotic expansions of equations (3)–(11) in $\epsilon \sim \delta b = \delta B_\perp / B_0$, with $\delta b / \delta b_{\max} = \delta b\, \beta^{1/2} \sim \mathcal{O}(1)$. These calculations formally justify some of the ideas presented in the main text; e.g., the flattening effects of the heat fluxes and the scaling of traveling wave damping. In addition, the theory allows the determination of the numerical value for the initial decay rate of a traveling wave (i.e., a_2 in equation (48)), and the form of the nonlinearity that arises from the $\mathcal{O}(\beta^{-1/2})$ spatially varying part of Δ near interruption (see equation (14)). In all cases, we consider a strictly one-dimensional wave, as in the main text.

Although an asymptotic expansion, as promised above, is in principle straightforward, in practice there are several issues that arise. Most importantly, our ordering scheme does not allow for a single wave equation that describes both the early evolution of a wave (i.e., when $1 + \beta\Delta/2 \sim \mathcal{O}(1)$) and the final approach to the interruption (when $1 + \beta\Delta/2 \ll 1$). This problem is related to Δ being spatially constant at lowest order because of the smoothing effect of the heat fluxes, even though the spatial variation in Δ plays a key role in forming square structures as the wave approaches the interruption limit. This motivates two separate expansions: the first is valid when $1 + \beta\Delta/2 \sim \mathcal{O}(1)$, the second is valid when $1 + \beta\Delta/2 \sim \mathcal{O}(\epsilon^2)$ (i.e., when the wave has already evolved to be close to the interruption limit).

These difficulties motivate our arrangement of this appendix as follows. We start in appendix A.2 by considering the double-adiabatic version of equations (3)–(7), with $q_\perp = q_\parallel = 0$. This leads to a simple nonlinear wave equation that is free from the issues mentioned in the previous paragraph, because there is large spatial variation in Δ. We then consider the initial wave evolution using the LF closure in appendix A.3, ordering $u_\perp / v_A \sim \delta B_\perp / B_0 \sim \mathcal{O}(\epsilon)$ and $1 + \beta\Delta/2 \sim \mathcal{O}(1)$, which can be used to find the initial damping rate of a traveling wave. Finally we derive an equation for waves as they get very close to the interruption limit itself in appendix A.4. This involves a spatially varying nonlinearity arising from both the field curvature and the spatial variation of Δ. Unfortunately, a closure problem prevents true asymptotic determination of the evolution of the spatial mean of Δ, although the expansion is still helpful for determining the residual spatial variation of Δ and formulating a simple nonlinear wave equation that describes the wave's approach to zig-zag field-line structures.

A.1. Nondimensionalized equations

For the sake of algebraic simplicity and to emphasize the appearances of β, throughout the following sections we work in dimensionless variables. These are chosen such that the Alfvén frequency and wavenumber are both unity:

$$\begin{gathered}
\boldsymbol{x} = k_\parallel^{-1}\bar{\boldsymbol{x}}, \quad t = \omega_A^{-1}\bar{t}, \quad \boldsymbol{u} = v_{A0}\bar{\boldsymbol{u}}, \quad \boldsymbol{B} = B_0\bar{\boldsymbol{B}}, \quad \rho = \rho_0\bar{\rho}, \quad p_\perp = p_0\bar{p}_\perp, \quad p_\parallel = p_0\bar{p}_\parallel, \\
q_\perp = c_s p_0 \bar{q}_\perp, \quad q_\parallel = c_s p_0 \bar{q}_\parallel, \quad \beta_0 \equiv \frac{8\pi p_0}{B_0^2}, \quad v_{A0} \equiv \frac{B_0}{\sqrt{4\pi\rho_0}}, \\
c_s^2 = 2\frac{p_0}{\rho_0} = \beta_0 v_{A0}^2, \quad \omega_A \equiv k_\parallel v_{A0}, \quad \nu_c = \omega_A \bar{\nu}_c, \quad \Delta \equiv \bar{p}_\perp - \bar{p}_\parallel.
\end{gathered} \tag{A.1}$$

Substituting these definitions into equations (3)–(7) and equations (10) and (11), one obtains

$$\partial_{\bar{t}}\bar{\rho} + \bar{\nabla}\cdot(\bar{\rho}\bar{\boldsymbol{u}}) = 0, \tag{A.2}$$

$$\bar{\rho}(\partial_{\bar{t}}\bar{\boldsymbol{u}} + \bar{\boldsymbol{u}}\cdot\bar{\nabla}\bar{\boldsymbol{u}}) = -\frac{\beta_0}{2}\bar{\nabla}\bar{p}_\perp - \bar{\nabla}\frac{\bar{B}^2}{2} + \bar{\nabla}\cdot\left[\hat{\boldsymbol{b}}\hat{\boldsymbol{b}}\left(\frac{\beta_0}{2}\Delta + \bar{B}^2\right)\right], \tag{A.3}$$

$$\partial_{\bar{t}}\bar{\boldsymbol{B}} + \bar{\boldsymbol{u}}\cdot\bar{\nabla}\bar{\boldsymbol{B}} = \bar{\boldsymbol{B}}\cdot\bar{\nabla}\bar{\boldsymbol{u}} - \bar{\boldsymbol{B}}\bar{\nabla}\cdot\bar{\boldsymbol{u}}, \tag{A.4}$$

$$\partial_{\bar{t}}\bar{p}_\perp + \bar{\nabla}\cdot(\bar{p}_\perp\bar{\boldsymbol{u}}) + \bar{p}_\perp\bar{\nabla}\cdot\bar{\boldsymbol{u}} + \beta_0^{1/2}[\bar{\nabla}\cdot(\bar{q}_\perp\hat{\boldsymbol{b}}) + \bar{q}_\perp\bar{\nabla}\cdot\hat{\boldsymbol{b}}] = \bar{p}_\perp\hat{\boldsymbol{b}}\hat{\boldsymbol{b}}:\bar{\nabla}\bar{\boldsymbol{u}} - \bar{\nu}_c\Delta, \tag{A.5}$$

$$\partial_{\bar{t}}\bar{p}_\parallel + \bar{\nabla}\cdot(\bar{p}_\parallel\bar{\boldsymbol{u}}) + \beta_0^{1/2}[\bar{\nabla}\cdot(\bar{q}_\parallel\hat{\boldsymbol{b}}) - 2\bar{q}_\perp\bar{\nabla}\cdot\hat{\boldsymbol{b}}] = -2\bar{p}_\parallel\hat{\boldsymbol{b}}\hat{\boldsymbol{b}}:\bar{\nabla}\bar{\boldsymbol{u}} + 2\bar{\nu}_c\Delta, \tag{A.6}$$

$$\bar{q}_\perp = -\sqrt{\frac{\bar{p}_\parallel}{\pi\bar{\rho}}}\frac{1}{|k_\parallel| + \bar{\nu}_c(\beta_0\pi\bar{p}_\parallel/\bar{\rho})^{-1/2}}\left[\bar{\rho}\bar{\nabla}_\parallel\left(\frac{\bar{p}_\perp}{\bar{\rho}}\right) - \bar{p}_\perp\left(1 - \frac{\bar{p}_\perp}{\bar{p}_\parallel}\right)\frac{\bar{\nabla}_\parallel\bar{B}}{\bar{B}}\right], \tag{A.7}$$

$$\bar{q}_\parallel = -2\sqrt{\frac{\bar{p}_\parallel}{\pi\bar{\rho}}}\frac{1}{|k_\parallel| + (3\pi/2 - 4)\bar{\nu}_c(\beta_0\pi\bar{p}_\parallel/\bar{\rho})^{-1/2}}\bar{\rho}\bar{\nabla}_\parallel\left(\frac{\bar{p}_\parallel}{\bar{\rho}}\right), \tag{A.8}$$

where $\bar{B}^2 \equiv \bar{\boldsymbol{B}}\cdot\bar{\boldsymbol{B}}$, $\hat{\boldsymbol{b}} \equiv \bar{\boldsymbol{B}}/B$, and $\bar{\nabla} = k_\parallel^{-1}\nabla$. The bars on variables are henceforth suppressed for clarity.

We shall carry out all calculations in 1D in z (as in the main text) with the imposed background magnetic field $\boldsymbol{B} = B_0\hat{\boldsymbol{z}}$ (this is $\bar{\boldsymbol{B}} = 1\hat{\boldsymbol{z}}$ in the dimensionless variables (A.1)). Magnetic-field perturbations $\delta\boldsymbol{B}_\perp$ are taken to be in the $\hat{\boldsymbol{x}}$ direction, which implies that $\hat{\boldsymbol{y}}$ directed vector components of $\boldsymbol{u}$ and $\boldsymbol{B}$ are identically zero. Note that because $B_0 = 1$, the B_x used throughout the following sections is the same as the δb used in the main text (e.g., equation (20)). Our expansion is carried out in $\epsilon \sim B_x$ with $B_x\beta_0^{1/2} \sim \epsilon^0$, which implies $\beta_0 \sim \epsilon^{-2}$. With such a scaling ($\beta_0 \gg 1$), it is immediately apparent that the pressure terms (in the momentum equation (A.3)) and the heat-flux terms (in the pressure equations (A.5)–(A.6)) dominate, since all space and time derivatives for an Alfvén wave are $\sim\mathcal{O}(1)$.

Throughout the following sections we also define the spatial average

$$\langle f\rangle \equiv \frac{1}{2\pi}\int \mathrm{d}z\, f, \tag{A.9}$$

and the spatially varying part of a quantity

$$\tilde{f} \equiv f - \langle f\rangle. \tag{A.10}$$

A.2. The double-adiabatic limit

It transpires that an asymptotic expansion of the double-adiabatic equations ($q_\perp = q_\parallel = 0$ in equations (A.2)–(A.6)) is significantly simpler than that with the Landau closure for the heat fluxes (equations (A.7) and (A.8)). This is because a nonlinearity with spatial dependence appears in the lowest-order wave equation, due to the spatial variation of Δ being comparable to its mean. For this reason we start by outlining the procedure for the double-adiabatic equations, even though the neglect of the heat fluxes is not a valid approximation in the high-β limit.

Our asymptotic ordering, motivated by our interest in solutions near the interruption limit with $\delta B_\perp/B_0 \sim \beta_0^{1/2}$, is

$$B_x = \epsilon B_{x1} + \epsilon^2 B_{x2} + \mathcal{O}(\epsilon^3), \tag{A.11}$$

$$u_x = \epsilon u_{x1} + \epsilon^2 u_{x2} + \mathcal{O}(\epsilon^3), \tag{A.12}$$

$$u_z = \epsilon^2 u_{z2} + \epsilon^3 u_{z3} + \mathcal{O}(\epsilon^4), \tag{A.13}$$

$$\rho = 1 + \epsilon^2\rho_2 + \epsilon^3\rho_3 + \mathcal{O}(\epsilon^4), \tag{A.14}$$

$$p_\perp = 1 + \epsilon^2 p_{\perp 2} + \epsilon^4 p_{\perp 4} + \mathcal{O}(\epsilon^4), \tag{A.15}$$

$$p_\parallel = 1 + \epsilon^2 p_{\parallel 2} + \epsilon^3 p_{\parallel 3}\mathcal{O}(\epsilon^4). \tag{A.16}$$

In addition, $B_z = 1$ due to $\nabla\cdot\boldsymbol{B} = 0$. This leads to

$$B = 1 + \frac{1}{2}\epsilon^2 B_{x1}^2 + \mathcal{O}(\epsilon^3), \tag{A.17}$$

$$\hat{b}_x = \epsilon B_{x1} + \epsilon^2 B_{x2} + \epsilon^3\left(B_{x3} - \frac{B_{x1}^3}{2}\right) + \mathcal{O}(\epsilon^4), \tag{A.18}$$

$$\hat{b}_z = 1 - \frac{1}{2}\epsilon^2 B_{x1}^2 + \mathcal{O}(\epsilon^3), \tag{A.19}$$

$$\Delta = \epsilon^2(p_{\perp 2} - p_{\parallel 2}) + \epsilon^3(p_{\perp 3} - p_{\parallel 3}) + \mathcal{O}(\epsilon^4) = \epsilon^2\Delta_2 + \epsilon^3\Delta_3 + \mathcal{O}(\epsilon^4). \tag{A.20}$$

We now insert the expansions (A.11)–(A.16) into equations (A.2)–(A.6) and expand in ϵ to obtain a wave equation for B_{x1}.

Order ϵ^0. There is only one contribution at $\mathcal{O}(\epsilon^0)$, which comes from the z component of equation (A.3),

$$-\frac{\beta_0}{2}\partial_z p_{\perp 2} + \frac{\beta_0}{2}\partial_z(\hat{b}_{z0}\hat{b}_{z0}p_{\perp 2} - p_{\parallel 2}) = -\frac{\beta_0}{2}p_{\parallel 2} = 0. \tag{A.21}$$

This says that $p_{\parallel 2}$ has no spatial variation, $\tilde{p}_{\parallel 2} = 0$, expressing the parallel pressure balance.

Order ϵ^1. The perpendicular velocity at $\mathcal{O}(\epsilon)$ satisfies

$$\partial_t u_{x1} = \partial_z\left[B_{x1}\left(1 + \frac{\beta_0}{2}\Delta_2\right)\right], \tag{A.22}$$

while the induction equation (A.4) is simply

$$\partial_t B_{x1} = \partial_z u_{x1}. \tag{A.23}$$

The parallel momentum equation (A.3) again gives $\partial_z p_{\parallel 3} = 0$ and there is no contribution from the continuity (A.2) or pressure equations (A.5) and (A.6). From equation (A.22), it is clear that we need an expression for Δ_2, and, for the system to be closed, this must depend only on u_{x1} and B_{x1}.

Order ϵ^2. To calculate Δ_2, we require only $p_{\perp 2}$ and $p_{\parallel 2}$, so may ignore the momentum (A.3) and induction (A.4) equations at this order. Noting that $\hat{\boldsymbol{b}}\hat{\boldsymbol{b}} : \nabla \boldsymbol{u} = \hat{b}_z\hat{b}_z\partial_z u_z + \hat{b}_x\hat{b}_z\partial_z u_x$, the perpendicular and parallel pressure equations (A.5) and (A.6) become

$$\partial_t p_{\perp 2} + \partial_z u_{z2} = B_{x1}\partial_z u_{x1}, \tag{A.24}$$

$$\partial_t p_{\parallel 2} + 3\partial_z u_{z2} = -2B_{x1}\partial_z u_{x1}. \tag{A.25}$$

From the $\mathcal{O}(\epsilon^0)$ parallel momentum equation (A.21), we know that $\tilde{p}_{\parallel 2} = 0$, so, using $\langle \partial_z u_{z2}\rangle = 0$, we obtain

$$\partial_z u_{z2} = -\frac{2}{3}\widetilde{B_{x1}\partial_z u_{x1}}. \tag{A.26}$$

Combining this with the perpendicular pressure equation (A.24) and assuming $\Delta_2(t=0) = 0$, we find

$$\Delta_2 = \frac{5}{6}(B_{x1}(t)^2 - B_{x1}(0)^2) + \frac{2}{3}\langle B_{x1}(t)^2 - B_{x1}(0)^2\rangle. \tag{A.27}$$

Inserting equation (A.27) into equation (A.22) and using equation (A.23), we obtain a closed wave equation for B_{x1}:

$$\partial_t^2 B_{x1} = \partial_z^2\left[B_{x1}\left(1 + \frac{5}{12}\beta_0[B_{x1}^2]_0^t + \frac{1}{3}\beta_0\langle[B_{x1}^2]_0^t\rangle\right)\right], \tag{A.28}$$

where $[f]_0^t \equiv f(t) - f(0)$.

Because there is a strong spatially varying nonlinearity arising from the $[B_{x1}^2]_0^t$ term, equation (A.28) can represent solutions of the full double-adiabatic equations (A.2)–(A.6) both initially and near the interruption limit. As illustrated in figure A1, where we compare solutions of the full double-adiabatic model with those of the wave equation (A.28), this nonlinearity causes a significant change in the shape of the wave as it evolves. As should be expected because of the different spatial form of the nonlinearity, the nonlinear evolution at the interruption limit is quite different from the evolution when heat fluxes are included (see figure 6). In particular, since the heat fluxes no longer act to spatially smooth the pressure anisotropy, B^2 can vary in space even with the anisotropy everywhere close to the firehose limit. The model equation does a reasonably good job at capturing the main qualitative features of interruption and is nearly perfect for the initial wave interruption (compare dashed lines in each panel). It is worth noting that the relative spatial variation of $p_\perp$ compared to its mean (and the lack of variation in $p_\parallel$) can also be obtained by considering the compressible part of the CGL equations as a forced oscillator system (see Squire and Quataert 2017), and this agrees with numerical solutions of the full equations (A.2)–(A.6).

A.3. LF closure: initial evolution

Here, we repeat the calculation of the previous section but include the LF prescription (A.7) and (A.8) for the heat fluxes. These act to smooth pressure perturbations on the sound-crossing timescale, leading to a pressure perturbation that is constant in space to lowest order, i.e., $\widetilde{\Delta_2} = 0$. The resulting equation for B_x is thus not accurate when $1 + \beta_0\Delta/2 \ll 1$ (i.e., in the approach to the interruption limit), because there is no spatially local nonlinearity to steepen the wave into zig-zag structures during the slow dynamics when $\beta_0\Delta/2 \approx -1$. As well as motivating the use of a second expansion with $1 + \beta_0\Delta/2$ ordered small (this is done in the appendix A.4), the expansion presented here is used to calculate the damping rate of a traveling wave due to the spatial correlation of $B^{-1}\mathrm{d}B/\mathrm{d}t$ and Δ, as used in the arguments in section 5.2. This is nonzero only at order $\mathcal{O}(\epsilon^3)$, because $\widetilde{\Delta_2} = 0$.

We use the same ordering as the previous section, equations (A.11)–(A.16). To lowest order, the heat fluxes (A.7) and (A.8) simplify significantly to

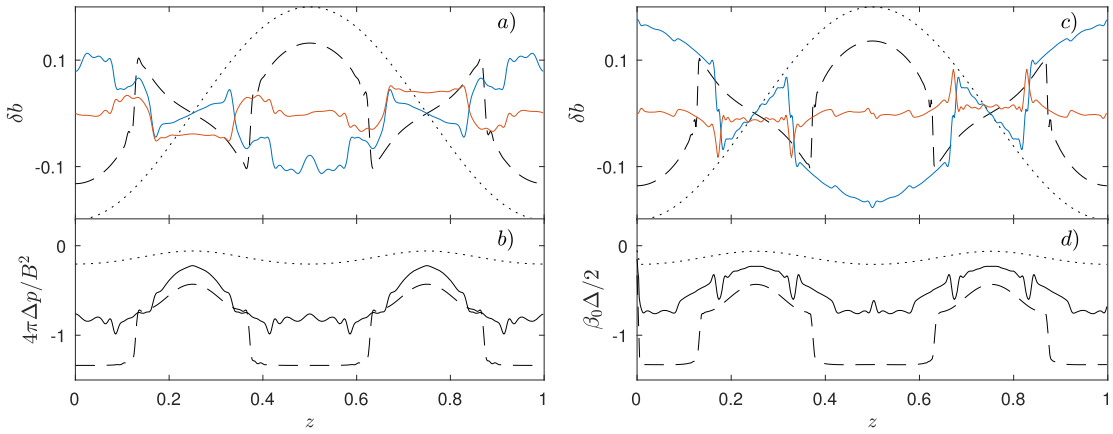

Figure A1. Evolution of an initial magnetic perturbation (standing wave; initial conditions (21)) within the CGL model, at $\beta_0 = 100$ with $\delta b\,(t=0) = B_{x1}(0) = 0.2$. Panels (a) and (b) show the solutions of the full CGL model (equations (A.2)–(A.6)), while panels (c) and (d) shows the solution to the asymptotic wave equation (A.28). The top panels show δb at $t = 0$ (black dotted line), δb at $t = 0.25\tau_A$ (black dashed line), δb at $t = \tau_A$ (blue solid line), and u_x/v_A at $t = \tau_A$ (red solid line). The bottom panels show the pressure anisotropy parameter, $4\pi\Delta p/B^2$ in panel (b) and $\beta_0\Delta/2$ in panel (d) (these are -1 at the firehose limit in each case), at $t = 0.05\tau_A$ (dotted line), $t = 0.25\tau_A$ (dashed line), and $t = \tau_A$ (solid line). Note the differences between these solutions and the evolution of a wave with the Landau fluid prescription for the heat fluxes (e.g., figure 6), which should be expected due to the different form of the spatially varying nonlinearity.

$$q_\perp = \epsilon^2\,\pi^{-1/2}\frac{\partial_z}{|k_z|}(p_{\perp 2} - \rho_2) + \epsilon^3\,\pi^{-1/2}\,\frac{\partial_z}{|k_z|}(p_{\perp 3} - \rho_3) + \mathcal{O}(\epsilon^4), \tag{A.29}$$

$$q_\parallel = \epsilon^2\,2\pi^{-1/2}\frac{\partial_z}{|k_z|}(p_{\parallel 2} - \rho_2) + \epsilon^3\,2\pi^{-1/2}\,\frac{\partial_z}{|k_z|}(p_{\parallel 3} - \rho_3) + \mathcal{O}(\epsilon^4), \tag{A.30}$$

while $\nabla\cdot(q_\perp\hat{\boldsymbol{b}}) = \epsilon^2\partial_z q_{\perp,2} + \epsilon^3\partial_z q_{\perp,3} + \mathcal{O}(\epsilon^4)$ (and similarly for $q_\parallel$). These simplifications are tantamount to stating that the heat flows along the mean field (i.e., along $\hat{\boldsymbol{z}}$) at the lowest two orders. As we did for the double-adiabatic calculation in appendix A.2, let us go through each order of the expanded equations.

Order ϵ^0. This is unchanged from the double-adiabatic calculation, giving $\tilde{p}_{\parallel 2} = 0$ due to parallel pressure balance (see equation (A.21)).

Order ϵ^1. The perpendicular momentum equation (A.3) and induction equation (A.4) at this order remain unchanged compared to the double-adiabatic result (equations (A.22) and (A.23)). The parallel momentum equation (A.3) is also unchanged, giving

$$\tilde{p}_{\parallel 3} = 0. \tag{A.31}$$

However, due to the $\beta_0^{1/2}$ terms in equations (A.5) and (A.6), there is now a contribution at $\mathcal{O}(\epsilon)$ in the pressure equations,

$$\pi^{-1/2}\beta_0^{1/2}|k_z|(p_{\perp 2} - \rho_2) = 0, \tag{A.32}$$

$$2\pi^{-1/2}\beta_0^{1/2}|k_z|(p_{\parallel 2} - \rho_2) = 0, \tag{A.33}$$

where we have used $\partial_z^2/|k_z| = -|k_z|$ to simplify the nonlocal diffusion operators. Combined with $\partial_z p_{\parallel 2} = 0$ and with the $\mathcal{O}(\epsilon^2)$ continuity equation $\partial_t\rho_2 + \partial_z u_{z2} = 0$, equations (A.32) and (A.33) imply

$$\tilde{p}_{\perp 2} = \widetilde{\rho_2} = \widetilde{u_{z2}} = 0, \tag{A.34}$$

meaning that the pressure anisotropy is spatially constant to lowest order, $\widetilde{\Delta_2} = 0$. This contrasts with the double-adiabatic result (A.24) and formally justifies the discussions in the main text regarding the smoothing effects of the heat fluxes.

Order ϵ^2. Again, the momentum and induction equations (A.3) and (A.4) are not useful for our purposes at order $\mathcal{O}(\epsilon^2)$, so we consider only the pressure equations (A.5) and (A.6). These are

$$\partial_t p_{\perp 2} + \partial_z u_{z2} + \pi^{-1/2}\beta_0^{1/2}|k_z|(p_{\perp 3} - \rho_3) = B_{x1}\partial_z u_{x1}, \tag{A.35}$$

$$\partial_t p_{\parallel 2} + 3\,\partial_z u_{z2} + 2\pi^{-1/2}\beta_0^{1/2}|k_z|(p_{\parallel 3} - \rho_3) = -2B_{x1}\partial_z u_{x1}, \tag{A.36}$$

where we also know from order $\mathcal{O}(\epsilon)$ that $\partial_z u_{z2} = 0$, $\langle p_{\parallel 2}\rangle = p_{\parallel 2}$, and $\langle p_{\perp 2}\rangle = p_{\perp 2}$ (equation (A.34)). Averaging equations (A.35) and (A.36) and solving the resulting equation assuming $\Delta_2(t=0) = 0$ gives, therefore,

$$\Delta_2 = \frac{3}{2}\langle B_{x1}(t)^2 - B_{x1}(0)^2\rangle. \tag{A.37}$$

This leads to the following wave equation for B_{x1}:

$$\partial_t^2 B_{x1} = \partial_z^2\left[B_{x1}\left(1 + \frac{3}{4}\beta_0\langle[B_{x1}^2]_0^t\rangle\right)\right], \tag{A.38}$$

which is the same as equation (20) in the main text, with equation (14) for the anisotropy and neglecting the δb^3 terms that arise from the magnetic curvature. The problems with using equation (A.38) to describe the wave near $1 + \beta_0\Delta/2 = 0$ are immediately apparent: no matter how small one takes ϵ, there will always be a time for which the higher-order contributions from the magnetic curvature and spatial variation of Δ play an important dynamical role. Indeed, numerical solutions to equation (A.38) stay perfectly sinusoidal until $1 + \beta_0\Delta/2 < 0$, at which point small-scale (firehose) fluctuations grow rapidly. There is no tendency for the wave to become square. This issue will be resolved in appendix A.4 through the use of a different ordering scheme.

A.3.1. Wave damping. A traveling wave, which satisfies $\langle B_x^2\rangle = \text{const.}$, propagates linearly, with no nonlinear modification, under equation (A.38). Although a continuation of the expansion to higher order is not very useful under this ordering, one can obtain an estimate of the lowest-order contribution to the damping of the wave energy into thermal energy that occurs for a traveling wave due to the spatial dependence of $p_{\perp 3}$. In the dimensionless variables (A.1), the kinetic-energy evolution equation (17) is

$$\partial_t\langle\rho u^2\rangle + \partial_t\langle B^2\rangle = \beta_0\langle p_\parallel\nabla\cdot\boldsymbol{u}\rangle - \beta_0\left\langle\Delta\frac{1}{B}\frac{\mathrm{d}B}{\mathrm{d}t}\right\rangle. \tag{A.39}$$

The right-hand side of this equation includes compressional and pressure-anisotropy heating, which can cause the transfer of mechanical wave energy into thermal energy. Using the variable expansions (A.11)–(A.16) and equation (A.34), equation (A.39) becomes

$$\begin{aligned}\partial_t\langle\rho u^2 + B^2\rangle = &-\epsilon^2\frac{\beta_0}{2}\Delta_2\partial_t\langle B_{x1}^2\rangle\\ &-\epsilon^3\beta_0\left[\Delta_2\partial_t\langle B_{x1}B_{x2}\rangle + \frac{1}{2}\langle\Delta_3\rangle\partial_t\langle B_{x1}^2\rangle + \frac{1}{2}\langle\widetilde{\Delta_3}\partial_t(B_{x1}^2)\rangle\right] + \mathcal{O}(\epsilon^4),\end{aligned} \tag{A.40}$$

where we have split Δ_3 into its mean and spatially varying parts, $\langle\Delta_3\rangle$ and $\widetilde{\Delta_3}$. The compressional term $\langle p_\parallel\nabla\cdot\boldsymbol{u}\rangle$ contributes only at order $\mathcal{O}(\epsilon^5)$ and higher, because $\partial_z p_{\parallel 2} = \partial_z p_{\parallel 3} = 0$ (equation (A.31)) and $\partial_z u_{z2} = 0$ (equation (A.34)). The $\mathcal{O}(\epsilon^2)$ term on the right-hand side of equation (A.40) is zero for a traveling wave, because $\partial_t\langle B_{x1}^2\rangle = 0$, and similarly for the second of the order $\mathcal{O}(\epsilon^3)$ terms. Since we are interested in the damping of a pure sine wave, we shall also ignore the first order $\mathcal{O}(\epsilon^3)$ term, which is related to the development of shape changes in B_x. This leaves us with $(\beta_0/2)\langle\widetilde{\Delta_3}\partial_t(B_{x1}^2)\rangle$ in the right-hand side of equation (A.40). This term describes how the average spatial correlation of Δ with $B^{-1}\mathrm{d}B/\mathrm{d}t$ causes a net damping, even when the averages of Δ and $B^{-1}\mathrm{d}B/\mathrm{d}t$ individually are each zero.

To calculate $\widetilde{\Delta_3}$, consider the spatially varying part of equations (A.35) and (A.36):

$$\pi^{-1/2}\beta_0^{1/2}|k_z|(p_{\perp 3} - \rho_3) = \widetilde{B_{x1}\partial_z u_{x1}}, \tag{A.41}$$

$$2\pi^{-1/2}\beta_0^{1/2}|k_z|(p_{\parallel 3} - \rho_3) = -2\widetilde{B_{x1}\partial_z u_{x1}}. \tag{A.42}$$

Using $\tilde{p}_{\parallel 3} = 0$ (equation (A.31)), solving for ρ_3, and inserting this solution into the $p_\perp$ equation gives

$$\tilde{p}_{\perp 3} = \widetilde{\Delta_3} = 2\sqrt{\frac{\pi}{\beta_0}}|k_z|^{-1}\widetilde{B_{x1}\partial_z u_{x1}}. \tag{A.43}$$

Therefore, the third-order wave-damping rate is

$$\partial_t\langle\rho u^2 + B^2\rangle = 2\sqrt{\pi}\beta_0^{1/2}\langle\partial_t(B_{x1}^2)\partial_t[|k_z|^{-1}(B_{x1}^2)]\rangle. \tag{A.44}$$

For a traveling sinusoidal wave $B_{x1} = \delta b\cos(z - t)$, by carrying out the spatial integrations using $|k_z|^{-1}B_{x1}^2 = \cos(2z - 2t)/4$, we find that the wave energy is damped at the rate $\partial_t E_{\text{wave}} = \sqrt{\pi^3}\beta_0^{1/2}\delta b^4$, or, restoring dimensions,

$$\partial_t E_{\text{wave}} = \frac{\sqrt{\pi}}{8}\omega_A\beta_0^{1/2}B_0^2\delta b^4 \tag{A.45}$$

in Gauss units per unit length. This provides more formal justification (and the numerical coefficient) for the wave-damping rate (equation (48)) used to derive kinetic and magnetic energy damping rates (equation (52)) in section 5.2.

A.4. LF closure: approach to wave interruption

In this section, we derive a nonlinear wave equation to describe the final approach to wave interruption, which is not captured correctly by equation (A.38) because it lacks a spatially dependent nonlinearity. Specifically, the distance from marginality, $1 + \beta_0\Delta/2$, is ordered as $\mathcal{O}(\epsilon^2)$ even though 1 and $\beta_0\Delta/2$ are each $\mathcal{O}(\epsilon^0)$. Since the previous expansion yielded the result that $\widetilde{\Delta} = 0$ to lowest order (equation (A.34)), this may be considered as a re-ordering of the equations, which becomes valid when equation (A.38) loses its validity because $1 + \beta_0\Delta/2 \ll 1$. Under the assumption of small $1 + \beta_0\Delta/2$, we are also forced to assume $u_x \sim \epsilon B_x$ and $\partial_t \sim \epsilon$, as should be expected[30]. The resulting wave equation (A.60) contains spatially dependent nonlinearities from both the magnetic curvature and the spatial variation of Δ. It thus contains the terms necessary to reproduce zig-zag field-line structures seen in solutions of the full LF equations (e.g., figure 6).

Given these considerations, our asymptotic ordering is modified from equations (A.11)–(A.16) as follows:

$$\partial_t f \sim \epsilon\, f, \tag{A.46}$$

$$\Delta = -\frac{2}{\beta_0} + \epsilon^3\Delta_3 + \epsilon^4\Delta_4 + \mathcal{O}(\epsilon^5), \tag{A.47}$$

$$u_x = \epsilon^2 u_{x2} + \epsilon^3 u_{x3} + \mathcal{O}(\epsilon^4), \tag{A.48}$$

$$u_z = \epsilon^3 u_{z3} + \epsilon^4 u_{z4} + \mathcal{O}(\epsilon^5), \tag{A.49}$$

where f represents any variable (the change in the ordering of u_z stems from the time derivative in the continuity equation (A.2)). Before embarking on an order-by-order expansion, we can simplify our task significantly by noting that only every second term in each field expansion need be considered, viz., $B_x = \epsilon B_{x1} + \epsilon^3 B_{x3} + \mathcal{O}(\epsilon^5)$, $p_\perp = 1 + \epsilon^2 p_{\perp 2} + \epsilon^4 p_{\perp 4} + \mathcal{O}(\epsilon^6)$, $\Delta = -2/\beta_0 + \epsilon^4\Delta_4 + \mathcal{O}(\epsilon^6)$ etc, for all fields. This is justified by the fact that in all equations, the expressions for order-n quantities depend only on order-$(n+2)$ quantities, because the terms that relate to modifications in B all contain[31] B_x^2. We must still work through all orders of the equations in ϵ, since some fields contain even powers of ϵ (e.g., u_x, Δ) while others contain odd powers (e.g., B_x, u_z).

Order ϵ^0. This is unchanged from appendices A.2 and A.3, giving $\tilde{p}_{\parallel 2} = 0$ due to parallel pressure balance. Because $\Delta_2 = -2/\beta_0$ (equation (A.47)), this also implies $\tilde{p}_{\perp 2} = 0$.

Order ϵ^1. This remains unchanged from appendix A.3, giving $\tilde{p}_{\perp 2} = \widetilde{\rho_2} = 0$ due to the heat fluxes in the pressure equations (A.5) and (A.6). From the $\mathcal{O}(\epsilon^3)$ part of the continuity equation (A.2), $\rho_2 = 0$ implies $\widetilde{u_{z3}} = u_{z3} = 0$.

Order ϵ^2. The parallel momentum equation (A.3) is

$$-\frac{\beta_0}{2}\partial_z p_{\perp 4} - \frac{1}{2}\partial_z B_{x1}^2 + \frac{\beta_0}{2}\partial_z(\hat{b}_z\hat{b}_z\Delta) = 0. \tag{A.50}$$

Expanding $\hat{b}_z$ and using $\Delta = -2/\beta_0 + \mathcal{O}(\epsilon^4)$, this leads to

$$\partial_z p_{\parallel 4} = \frac{1}{\beta_0}\partial_z B_{x1}^2. \tag{A.51}$$

The perpendicular induction equation appears at this order, giving

$$\partial_t B_{x1} = \partial_z u_{x2} \tag{A.52}$$

as expected.

Order ϵ^3. The perpendicular momentum equation, which forms the basis for our wave equation, appears at $\mathcal{O}(\epsilon^3)$ and reads

$$\partial_t u_{x2} = \partial_z\left(\frac{\beta_0}{2}\Delta_4 B_{x1} + B_{x1}^3\right), \tag{A.53}$$

where we have used $1 + \beta_0\Delta_2/2 = 0$ (equation (A.47)). The final term in equation (A.53) arises from the $\mathcal{O}(\epsilon^2)$ contributions to $\hat{b}_x\hat{b}_z$ from the variation in B, which is of the same origin as the $1/(1+\delta b^2)$ in equation (20). To continue, we need an expression for $\beta_0\Delta_4/2$ in equation (A.53), which requires the pressure equations (A.5) and (A.6).

Because we assume $\Delta_2 = -2/\beta_0$, the second-order pressures are constant, $\partial_t p_{\perp 2} = \partial_t p_{\parallel 2} = 0$, and the time derivatives first occur in the pressure equations (A.5) and (A.6) at order $\mathcal{O}(\epsilon^5)$. Noting that $\partial_z u_{z3} = 0$, the

[30] The reader may notice that the spatially varying part of Δ was $\mathcal{O}(\epsilon^3)$ in the expansion of the previous section (equation (A.43)), whereas here, by assuming $1 + \beta_0\Delta/2 \sim \epsilon^2$ we are effectively ordering it to be $\mathcal{O}(\epsilon^4)$. This apparent discrepancy is resolved by noting that as the solutions of equation (A.38) evolve towards $1 + \beta_0\Delta/2 = 0$, the time derivatives (or equivalently u_x) also become one order smaller, meaning that the spatial variation of Δ is pushed into Δ_4 (recall that Δ_3 was determined by the current value of $\partial_t B_{x1}^2$, not its time history). Thus, the solutions of equation (A.38) will evolve into a regime where the expansion discussed in this section in valid.

[31] Note that this was not the case in the previous expansions (A.11)–(A.16) due to the $\beta_0^{1/2}$ coefficient of the heat fluxes in the pressure equations (A.5) and (A.6), whereas now there is an extra ϵ arising in the time derivatives of $p_{\perp,\parallel}$ in equations (A.5) and (A.6) that restores this property.

pressure equations (A.5) and (A.6) at $\mathcal{O}(\epsilon^3)$ are then,

$$\beta_0^{1/2}\partial_z q_{\perp 4} = B_{x1}\partial_z u_{x2}, \tag{A.54}$$

$$\beta_0^{1/2}\partial_z q_{\parallel 4} = -2B_{x1}\partial_z u_{x2}. \tag{A.55}$$

We can obtain useful information from both the spatial average and the spatially varying part of equations (A.54) and (A.55).

A spatial average of equations (A.54) and (A.55) leads to

$$\partial_t \langle B_{x1}^2\rangle = 0 + \mathcal{O}(\epsilon^5), \tag{A.56}$$

which implies that $\partial_t \langle B_{x1}^2\rangle \sim \epsilon^2 \partial_t B_{x1}^2$, i.e., that the spatial average of B_{x1}^2 varies in time more slowly than B_{x1}^2 itself. This can occur, for instance, if B_{x1} is increasing in some region and decreasing in another, as could occur in the approach to a square wave.

The spatially varying part of equations (A.54) and (A.55) can be used to solve for $\widetilde{\Delta_4}$. We first require the heat fluxes (A.7) and (A.8), which are[32],

$$q_{\perp 4} = -\pi^{-1/2}\frac{\partial_z}{|k_z|}(p_{\perp 4} - \rho_4) + \pi^{-1/2}\beta_0^{-1}\frac{\partial_z}{|k_z|}B_{x1}^2, \tag{A.57}$$

$$q_{\parallel 4} = -2\pi^{-1/2}\frac{\partial_z}{|k_z|}(p_{\parallel 4} - \rho_4). \tag{A.58}$$

Inserting equations (A.57) and (A.58) into equations (A.54) and (A.55), using $\partial_z p_{\parallel 4} = \beta_0^{-1}\partial_z B_{x1}^2$ (equation (A.51)), then solving for ρ_4 and inserting this into the $p_{\perp 4}$ equation yields

$$|k_z|\Delta_4 = 2\sqrt{\pi}\beta_0^{-1/2}\partial_t\widetilde{B_{x1}^2} + \beta_0^{-1}|k_z|\widetilde{B_{x1}^2}. \tag{A.59}$$

This may be inserted into equations (A.52) and (A.53) to yield the nonlinear wave equation

$$\partial_t^2 B_x = \frac{\beta_0}{2}\langle\Delta_4\rangle\partial_z^2 B_{x1} + \frac{3}{2}\partial_z^2(B_{x1}^3) - \frac{1}{2}\langle B_{x1}^2\rangle\partial_z^2 B_{x1} + \sqrt{\pi}\beta_0^{1/2}\partial_z^2[B_{x1}\partial_t(|k_z|^{-1}B_{x1}^2)]. \tag{A.60}$$

Unfortunately, equation (A.60) still contains the undetermined mean pressure anisotropy $\langle\Delta_4\rangle$. While one can, in principle, solve for $\langle\Delta_4\rangle$ by considering the mean part of the pressure equations (A.5) and (A.6), the result contains B_{x3}, so $\langle\Delta_4\rangle$ remains unknown. Of course, any attempt to subsequently solve for B_{x3} generates dependence on Δ_6, leading to a standard closure problem. Despite this issue, equation (A.60) remains useful for a number of reasons. First, the spatially dependent nonlinearities are interesting: because of the time derivative in $\widetilde{\Delta_4}$ (the first term on the right-hand side of equation (A.59)), this term has a diffusive effect in equation (A.60), and it can dissipate wave energy into thermal energy. This is not the case for the the $\partial_z^2(B_{x1}^3)$ nonlinearity in equation (A.60), which arises from spatial variation in the field strength. Secondly, the exact form of $\langle\Delta_4\rangle$ plays only a minor role, because it is simply a spatially constant number that must decrease as B^2 decreases (since $\partial_t\Delta \sim \partial_t B^2$). Indeed, it has only one property that is key to the dynamics described by equation (A.60)—it must be able to approach 0 and become negative, so as to slow the linear dynamics and allow the nonlinear terms to dominate. Corrections (at order ϵ^4) to the exact point at which this zero crossing occurs will presumably not affect the dynamics of the wave strongly. For the purposes of exploring solutions to equation (A.60) numerically (see figure A2), we thus make the simple ansatz

$$\langle\Delta_4\rangle = \langle\Delta_4(0)\rangle + \frac{3}{4}\beta_0\langle(B_{x1}^2 - B_{x1}(0)^2)\rangle. \tag{A.61}$$

Here $\langle\Delta_4(0)\rangle$ is the initial anisotropy, which would arise through dynamics that satisfy equation (A.38). The second term comes from the form of Δ_2 in equation (A.38) (but it is now of order $\mathcal{O}(\epsilon^4)$ for the same reason that we obtained equation (A.56)), which is a clear choice that satisfies the requirements discussed above.

As shown in figure A2, the nonlinear wave equation (A.60) with $\langle\Delta_4\rangle$ given by equation (A.61) has solutions that are pleasingly similar to those of the full LF equations (see figures 6–8): we see a clear approach to zig-zag field lines, a much faster decay of the velocity in comparison to magnetic field, the eruption of small-scale firehose modes at early times for a standing wave, and the slowing of a traveling wave as the anisotropy approaches the firehose limit.

Appendix B. Asymptotic wave equations—Braginskii limit

In this appendix, we derive asymptotic wave equations in the Braginskii limit, with $\omega_A \ll \nu_c$. As expected, this calculation is significantly simpler than the collisionless cases discussed appendix A. We find two regimes, with

[32] Note that the $p_\perp(1 - p_\perp/p_\parallel)\nabla_\parallel B/B$ part of $q_\perp$ in equation (A.7) has made an appearance at this order.

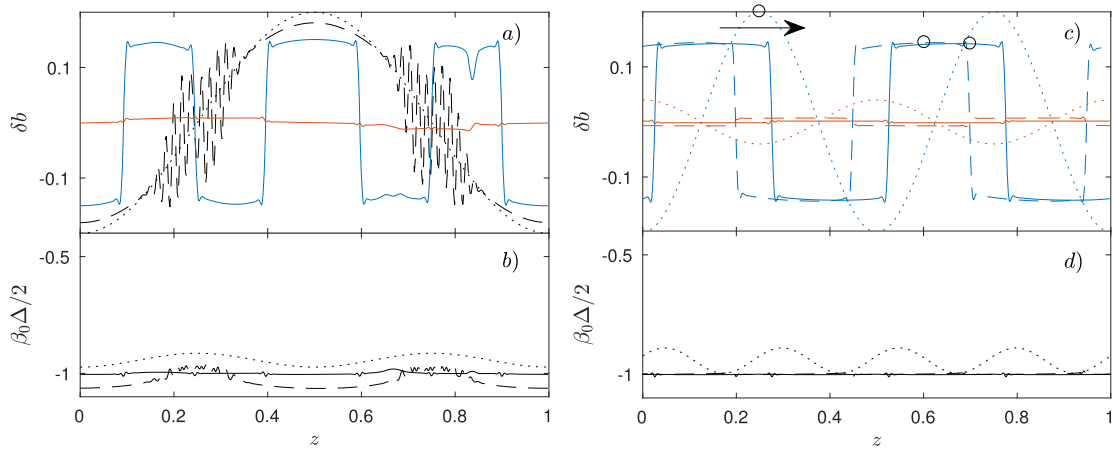

Figure A2. Solution of the 'approach-to-interruption' wave equation (A.60), using the closure (A.61) for $\langle\Delta_4\rangle$, at $\beta_0 = 100$, $B_{x1} = -0.2\cos(2\pi z)$, $\langle\Delta_4(0)\rangle = -0.2^2$. Panels (a) and (b) show a standing wave starting from $u_{x2} = 0.2^2\sin(2\pi z)$. Panel (a) shows the field and flow evolution: B_{x1} at $t = 0$ (black dotted line), B_{x1} at $t = 0.3\tau_A$ (black dashed line), B_{x1} at $t = 3\tau_A$ (blue solid line), and u_{x2} at $t = 3\tau_A$ (red solid line). Panel (b) shows the pressure anisotropy, $\beta_0\Delta/2$, at $t = 0$ (dotted line), $t = 0.3\tau_A$ (dashed line), and $t = 3\tau_A$ (solid line). Panels (c) and (d) illustrate a traveling wave near interruption, starting from $u_{x2} = 0.2^2\cos(2\pi z)$. Panel (c) shows B_{x1} (blue lines) and u_{x2} (red lines) at $t = 0$ (dotted lines), $t = 3\tau_A$ (dashed lines), $t = 6\tau_A$ (solid lines), with the circled points showing the same point of the wave as it travels to the right. Panel (d) shows the pressure anisotropy at $t = 0$ (dotted line), $t = 3\tau_A$ (dashed line), $t = 6\tau_A$ (solid line). The dynamics are very similar to those seen in the solution of the full Landau fluid equations (A.2)–(A.8) (see figures 6–8).

the dynamics controlled by the parameter $\bar{\nu}_c/\beta_0^{1/2}$ (recall that $\bar{\nu}_c \equiv \nu_c/\omega_A$)[33], which determines whether the effect of heat fluxes is important for the spatial form of the pressure anisotropy. In the first regime, when $\bar{\nu}_c \gg \beta_0^{1/2}$, one recovers the Braginskii wave equation discussed in the main text (equation (20) with the closure (9)); in the second, when $\bar{\nu}_c \sim \beta_0^{1/2}$ (or $\bar{\nu}_c < \beta_0^{1/2}$), one finds an anisotropy that becomes smoother in space as $\beta_0^{1/2}/\bar{\nu}_c$ increases. The magnetic field and velocity dynamics in each regime are generally similar to each other, heuristically vindicating the neglect of the heat fluxes in the main text. Throughout this appendix, we use the definitions and dimensionless equations described in appendix A.1.

B.1. General considerations and ordering

For consistency with the collisionless calculations in appendix A, we again use $\epsilon \sim B_x$, with $\delta b/\delta b_{\max} = \delta b\ \beta^{1/2}/\bar{\nu}_c^{1/2} \sim \mathcal{O}(1)$, which implies

$$\bar{\nu}_c \sim \epsilon^2\beta_0. \tag{B.1}$$

Combined with the fundamental requirement for the Braginskii approximation to be valid $\bar{\nu}_c \gg 1$, equation (B.1) suggests $\beta_0 \sim \mathcal{O}(\epsilon^{-3})$ or larger. This contrasts the collisionless ordering scheme used in appendix A, where we took $\beta_0 \sim \mathcal{O}(\epsilon^{-2})$. Throughout this section, we shall take $u_x \sim B_x \sim \mathcal{O}(\epsilon)$, as in appendices A.2 and A.3.

To constrain further the ordering of $\bar{\nu}_c$ and β_0 individually, let us consider the basic scaling of the heat fluxes in the Braginskii regime. Ignoring—for reasons that will become clear momentarily—the effect of the collisionality on the heat fluxes, let us consider the pressure equations (A.5) and (A.6). Since $\bar{\nu}_c$ must dominate over ∂_t, these equations, to lowest order, will be dominated by the terms $\beta_0^{1/2}\partial_z q_{\perp,\parallel}$, $B_x\partial_z u_x$, and $\bar{\nu}_c\Delta$. The balance between the collisions and heat fluxes is thus controlled by $\bar{\nu}_c/\beta_0^{1/2}$: if $\bar{\nu}_c \sim \beta_0^{1/2}$, the heat fluxes will enter at the same order as collisional isotropization in the pressure equations, while if $\bar{\nu}_c \gg \beta_0^{1/2}$, the heat fluxes will simply cause minor (higher-order) corrections to the spatial form of Δ. Our final equations will thus depend on the ordering of $\bar{\nu}_c/\beta_0^{1/2}$. This leads us to two natural choices: the 'high-collisionality regime,' $\nu_c \sim \mathcal{O}(\epsilon^4)$, $\beta_0 \sim \mathcal{O}(\epsilon^6)$, and and 'the moderate-collisionality regime,' $\nu_c \sim \mathcal{O}(\epsilon^2)$, $\beta_0 \sim \mathcal{O}(\epsilon^4)$[34].

In the discussion above, we neglected to mention collisional modifications to the heat fluxes (see equations (A.7) and (A.8)). For the sake of qualitative discussion, this neglect is admissible because collisions reduce the heat fluxes at the same point, $\bar{\nu}_c \sim \beta_0^{1/2}$, as the heat fluxes become subdominant to SA wave dynamics. This can be seen from the form of the LF heat fluxes, equations (A.7) and (A.8). Expanding these assuming small $p_\parallel$ and ρ perturbations, and ignoring numerical coefficients, one finds that the heat fluxes scale as

[33] Throughout this section, we retain the bar on the dimensionless $\bar{\nu}_c$ (defined in equation (A.1)), so as to emphasize this hidden factor of ω_A.

[34] Note that the regime in which the heat fluxes dominate, viz., $\bar{\nu}_c \ll \beta_0^{1/2}$, will turn out to be a subset of the moderate-collisionality regime in appendix B.3. There is thus no need to treat it separately.

$$q_{\perp,\parallel} \sim -\frac{1}{|k_\parallel| + \bar{\nu}_c \beta_0^{-1/2}} \nabla_\parallel \left(\frac{p_{\perp,\parallel}}{\rho} \right). \tag{B.2}$$

Combined with the discussion of the previous paragraph, this shows that in the limit $\bar{\nu}_c \gg \beta_0^{1/2}$, where the heat fluxes were not important in the pressure equation, the heat fluxes are even further reduced, making their neglect *more* valid than it would otherwise be. In contrast, when $\bar{\nu}_c \sim \beta_0^{1/2}$, the heat fluxes are only moderately affected by collisionality. Thus, the effect of collisionality on the heat fluxes is simply to improve the validity of the high-collisionality ordering, while changing the moderate-collisionality results by $\mathcal{O}(1)$ numerical factors.

B.2. High-collisionality limit

We now consider the high-collisionality ordering, which is

$$\begin{aligned} B_x \sim \mathcal{O}(\epsilon), \quad u_x \sim \mathcal{O}(\epsilon), \quad \bar{\nu}_c \sim \mathcal{O}(\epsilon^{-4}), \quad \beta_0 \sim \mathcal{O}(\epsilon^{-6}), \\ p_\perp \sim p_\parallel \sim \rho \sim 1 + \mathcal{O}(\epsilon^6), \quad u_z \sim \mathcal{O}(\epsilon^6), \end{aligned} \tag{B.3}$$

where the ordering of $p_{\perp,\parallel}$ is taken from the requirement that $\beta_0 \Delta \sim 1$ at the lowest order (or equivalently, from $B_x \partial_z u_x \sim \bar{\nu}_c \Delta$ in the pressure equations (A.5) and (A.6)).

The $\mathcal{O}(1)$ and $\mathcal{O}(\epsilon)$ equations under this ordering are effectively the same as in the collisionless calculations in appendices A.2 and A.3—the parallel momentum equation (A.3) at $\mathcal{O}(1)$ gives

$$\partial_z p_{\parallel 6} = 0, \tag{B.4}$$

while the perpendicular momentum equation (A.3) at $\mathcal{O}(\epsilon)$ gives

$$\partial_t u_{x1} = \partial_z \left[B_{x1} \left(1 + \frac{\beta_0}{2} \Delta_6 \right) \right]. \tag{B.5}$$

At order $\mathcal{O}(\epsilon^2)$, we need to consider only the pressure equations (A.5) and (A.6), which both give the same result,

$$\bar{\nu}_c \Delta_6 = B_{x1} \partial_z u_{x1}. \tag{B.6}$$

Inserted into equation (B.5), equation (B.6) leads to the expected wave equation

$$\partial_t^2 B_{x1} = \partial_z^2 \left[B_{x1} \left(1 + \frac{\beta_0}{4\bar{\nu}_c} \partial_t B_{x1}^2 \right) \right], \tag{B.7}$$

which, as expected, is identical to the wave equation studied in the text with a Braginskii closure, viz., equation (20) with the closure (9) (aside from the neglect of the $\mathcal{O}(\epsilon^3)$ correction $1/(1 + \delta b^2)$ due to field-strength variation in space).

B.3. Moderate-collisionality limit

As discussed in appendix B.1, the most natural moderate-collisionality ($\bar{\nu}_c \sim \beta_0^{1/2}$) ordering is

$$\begin{aligned} B_x \sim \mathcal{O}(\epsilon), \quad u_x \sim \mathcal{O}(\epsilon), \quad \bar{\nu}_c \sim \mathcal{O}(\epsilon^{-2}), \quad \beta_0 \sim \mathcal{O}(\epsilon^{-4}), \\ p_\perp \sim p_\parallel \sim \rho \sim 1 + \mathcal{O}(\epsilon^4), \quad u_z \sim \mathcal{O}(\epsilon^4). \end{aligned} \tag{B.8}$$

Again, the parallel momentum equation (A.3) leads to $\partial_z p_{\parallel 4} = 0$ at order $\mathcal{O}(1)$, while the perpendicular momentum equation (A.3) at order $\mathcal{O}(\epsilon)$ leads to equation (B.5) with Δ_6 replaced by Δ_4. At order $\mathcal{O}(\epsilon^2)$, the pressure equations contain both the heat fluxes and collisional relaxation:

$$\sqrt{\frac{\beta_0}{\pi}} \frac{-\partial_z^2}{|k_z| + a_\perp \bar{\nu}_c \beta_0^{-1/2}} (p_{\perp 4} - \rho_4) = B_{x1} \partial_z u_{x1} - \bar{\nu}_c \Delta_4, \tag{B.9}$$

$$2\sqrt{\frac{\beta_0}{\pi}} \frac{-\partial_z^2}{|k_z| + a_\parallel \bar{\nu}_c \beta_0^{-1/2}} (p_{\parallel 4} - \rho_4) = -2 B_{x1} \partial_z u_{x1} + 2\bar{\nu}_c \Delta_4, \tag{B.10}$$

where the coefficients $a_\perp = \pi^{-1/2}$ and $a_\parallel = (3\pi/2 - 4)\pi^{-1/2}$ account for the difference between the collisional parallel and perpendicular heat fluxes (Catto and Simakov 2004). Averaging these equations over space gives

$$\langle \Delta_4 \rangle = \bar{\nu}_c^{-1} \langle B_{x1} \partial_z u_{x1} \rangle, \tag{B.11}$$

which is also true in the high-collisionality regime (for $\langle \Delta_6 \rangle$; see equation (B.6)). Noting that the parallel momentum equation (A.3) at $\mathcal{O}(1)$ gives $\tilde{p}_{\parallel 4} = 0$, and that $\widetilde{\Delta_4} = \tilde{p}_{\perp 4}$, we can solve for $|k_z| \rho_4$ to obtain

$$\widetilde{\Delta_4} = \frac{\Theta \widetilde{B_{x1} \partial_z u_{x1}}}{\beta_0^{1/2} \zeta(\bar{\nu}_c) |k_z| + \bar{\nu}_c \Theta}. \tag{B.12}$$

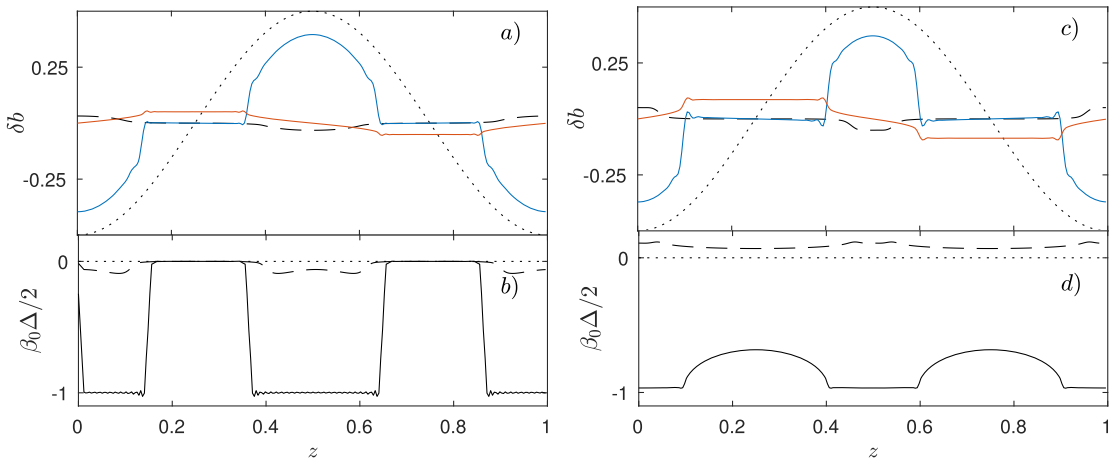

Figure B1. Evolution of the Braginskii wave equations (B.7) and (B.15), starting from a sinusoidal magnetic perturbation $B_{x1} = -0.5\cos(2\pi z)$. In the high-collisionality case shown in panels (a) and (b) we take $\bar{\nu}_c = 5^4 = 625$, $\beta = 5^6 = 15625$. In the moderate-collisionality case shown in panels (c) and (d) we take $\bar{\nu}_c = 5^2 = 25$, $\beta = 5^4 = 625$, such that $\bar{\nu}_c = \beta^{1/2}$ (we take $\zeta(\bar{\nu}_c) = (2\pi)^{-1/2}$ for simplicity). These parameters give the same interruption limit $\sqrt{\bar{\nu}_c/\beta} = 0.2$ in both cases. In the top panels, we show B_{x1} at $t = 0$ (dotted black line), B_{x1} at $t = 0.6\tau_A$ (solid blue line), u_{x1} at $t = 0.6\tau_A$ (solid red line), and B_{x1} at $t = 2\tau_A$ (dashed black line), which is after the wave has decayed to below the interruption limit. In the bottom panels, we show the pressure anisotropy, $\beta_0\Delta/2$, at the same times. Although the pressure anisotropy profiles are quite different in each case (compare panels (b) and (d)), the dynamics of the magnetic perturbation, including the time taken for the wave to decay, are similar. This is because the parts of the wave where $\Delta > -2/\beta$ have $B_{x1} = 0$ anyway (see discussion in text). Note that the sole difference between the calculation shown in figure 2 and that in panel (a) here is the $(1 + \delta b^2)^{-1}$ field nonlinearity term, which is not included here because it is at higher asymptotic order. Because the spatially varying nonlinearity due to Δp is larger than that due to δb^2, this term makes little difference to the dynamics.

Here the operator

$$\zeta(\bar{\nu}_c) = \frac{\pi^{-1/2}}{1 + |k_z|^{-1} a_\perp \bar{\nu}_c \beta_0^{-1/2}} \tag{B.13}$$

encapsulates the collisional quenching of the heat fluxes, and its effect changes from being a multiplication by $\pi^{-1/2}$ at $\bar{\nu}_c \ll \beta_0^{1/2}$, to an operator $\sim \beta^{1/2}\bar{\nu}_c^{-1}|k_z|$ at $\bar{\nu}_c \gg \beta_0^{1/2}$ (but in this limit, this term may be neglected in comparison to $\bar{\nu}_c\Theta$). Similarly, the operator

$$\Theta = 1 + \frac{|k_z| + a_\parallel \bar{\nu}_c \beta_0^{-1/2}}{|k_z| + a_\perp \bar{\nu}_c \beta_0^{-1/2}} \tag{B.14}$$

is effectively a multiplication by a factor between 2 (for $\bar{\nu}_c \ll \beta_0^{1/2}$) and ≈1.71 (for $\bar{\nu}_c \gg \beta_0^{1/2}$), which is necessary due to the numerical difference between perpendicular and parallel collisional heat fluxes.

Put together, equations (B.11) and (B.12), along with the perpendicular momentum equation (B.5), lead to the wave equation

$$\partial_t^2 B_{x1} = \partial_z^2\left[B_{x1} + B_{x1}\frac{\beta_0}{4\bar{\nu}_c}\left(\partial_t\langle B_{x1}^2\rangle + \frac{\bar{\nu}_c\Theta}{\zeta(\bar{\nu}_c)\beta_0^{1/2}|k_z| + \bar{\nu}_c\Theta}\widetilde{\partial_t B_{x1}^2}\right)\right]. \tag{B.15}$$

Evidently, this equation includes the high-collisionality limit, equation (B.7), when $\bar{\nu}_c \gg \beta_0^{1/2}$, viz., equation (B.7) is the $\bar{\nu}_c \gg \beta_0^{1/2}$ limit of equation (B.15). Equation (B.15) also captures a Braginskii version of the 'spatially constant-Δ' limit when $\bar{\nu}_c \ll \beta_0^{1/2}$.

Solutions to equations (B.7) and (B.15) (with $\beta_0^{1/2} = \bar{\nu}_c$, and taking $\Theta = 2$ and $\zeta(\bar{\nu}_c) = 1$ for simplicity) are illustrated in figure B1. It is interesting that even with $\beta_0^{1/2} = \bar{\nu}_c$ (figure B1(b)), when the heat fluxes significantly modify the pressure anisotropy, the dynamics are largely similar to the basic high-collisionality Braginskii limit discussed in the main text (figure B1(a); see also figure 2). The reason is for this is related to the nature of the Braginskii wave decay, as discussed in section 4. Effectively, the dynamics of the decaying wave separate into regions where $\Delta = -2/\beta_0$ and the field has curvature, and regions where $\Delta > -2/\beta_0$ and the field has no curvature (i.e., where the perturbed field δb is zero). The primary effect of the heat fluxes is thus to decrease the anisotropy where B_{x1} is already zero anyway, causing only small modifications to the dynamics of the standing wave. It is worth noting, however, that because Δp is smoothed more by the heat fluxes as $\nu_c/\beta_0^{1/2}$ decreases (i.e., the ratio of $\widetilde{\Delta_4}$ (B.12) to $\langle\Delta_4\rangle$ (B.11) decreases as $\nu_c/\beta_0^{1/2}$ decreases), the decay rate of a Braginskii traveling wave will be reduced by a factor between 1 and $\beta_0^{1/2}$ (compared to the estimate in the main text; see equation (37)), when the heat fluxes cause significant smoothing of the pressure anisotropy.

References

Achterberg A 1981 *Astron. Astrophys.* **98** 161
Balbus S A and Hawley J F 1998 *Rev. Mod. Phys.* **70** 1
Bale S D *et al* 2009 *Phys. Rev. Lett.* **103** 211101
Boldyrev S 2006 *Phys. Rev. Lett.* **96** 115002
Braginskii S I 1965 *Rev. Plasma Phys.* **1** 205
Bruno R and Carbone V 2013 *Living Rev. Solar Phys.* **10** 2
Catto P J and Simakov A N 2004 *Phys. Plasmas* **11** 90
Cesarsky C J and Kulsrud R M 1981 Cosmic-ray self confinement in the hot phase of the interstellar medium *Origin of Cosmic Rays* (*IAU Symp.* vol 94) ed G Setti p 251
Chandrasekhar S, Kaufman A N and Watson K M 1958 *Proc. R. Soc. Lond.* A **245** 435
Chew C F, Goldberger M L and Low F E 1956 *Proc. R. Soc. Lond.* A **236** 112
Cramer N F 2011 *The Physics of Alfvén Waves* (New York: Wiley)
Davidson R C and Völk H J 1968 *Phys. Fluids* **11** 2259
Del Zanna L, Velli M and Londrillo P 2001 *Astron. Astrophys.* **367** 705
Derby N F Jr 1978 *Astrophys. J.* **224** 1013
Eastwood J P *et al* 2005 *Space Sci. Rev.* **118** 41
Flå T, Mjølhus E and Wyller J 1989 *Phys. Scr.* **40** 219
Foote E A and Kulsrud R M 1979 *Astrophys. J.* **233** 302
Forest C B *et al* 2015 *J. Plasma Phys.* **81** 345810501
Galeev A A and Oraevskii V N 1963 *Sov. Phys.—Dokl.* **7** 988
Galtier S, Nazarenko S V, Newell A C and Pouquet A 2000 *J. Plasma Phys.* **63** 447
Gekelman W *et al* 2016 *Rev. Sci. Instrum.* **87** 025105
Gizon L, Cally P S and Leibacher J 2008 *Helioseismology, Asteroseismology, and MHD Connections* (Berlin: Springer)
Goldreich P and Sridhar S 1995 *Astrophys. J.* **438** 763
Goldstein M L 1978 *Astrophys. J.* **219** 700
Goswami P, Passot T and Sulem P L 2005 *Phys. Plasmas* **12** 102109
Hammett G W, Dorland W and Perkins F W 1992 *Phys. Fluids* B **4** 2052
Hammett G W and Perkins F W 1990 *Phys. Rev. Lett.* **64** 3019
Heidbrink W W 2008 *Phys. Plasmas* **15** 055501
Hellinger P and Matsumoto H 2000 *J. Geophys. Res.* **105** 10519
Hellinger P, Trávníček P, Kasper J C and Lazarus A J 2006 *Geophys. Res. Lett.* **33** L09101
Hellinger P and Trávníček P M 2008 *J. Geophys. Res.: Space* **113** A10109
Hellinger P and Trávníček P M 2015 *J. Plasma Phys.* **81** 305810103
Hollweg J V 1971a *J. Geophys. Res.* **76** 5155
Hollweg J V 1971b *Phys. Rev. Lett.* **27** 1349
Hunana P and Zank G P 2017 *Astrophys. J.* **839** 13
Klein K G and Howes G G 2015 *Phys. Plasmas* **22** 032903
Kulsrud R M 1978 *Astronomical Papers Dedicated to Bengt Stromgren* ed A Reiz and T Andersen pp 317–26
Kulsrud R M 1983 MHD description of plasma *Basic Plasma Physics: Selected Chapters, Handbook of Plasma Physics, Volume 1* ed R N Sagdeev and M N Rosenbluth (Amsterdam: North-Holland)
Kunz M W, Bogdanović T, Reynolds C S and Stone J M 2012 *Astrophys. J.* **754** 122
Kunz M W, Schekochihin A A, Chen C H K, Abel I G and Cowley S C 2015 *J. Plasma Phys.* **81** 325810501
Kunz M W, Schekochihin A A, Cowley S C, Binney J J and Sanders J S 2010 *Mon. Not. R. Astron. Soc.* **410** 2446
Kunz M W, Schekochihin A A and Stone J M 2014 *Phys. Rev. Lett.* **112** 205003
Lee M A and Völk H J 1973 *Astrophys. Space Sci.* **24** 31
Marsch E 2006 *Living Rev. Sol. Phys.* **3** 1
Matteini L, Landi S, Velli M, Zanna L D and Hellinger P 2010 *Geophys. Res. Lett.* **37** L20101
Medvedev M V and Diamond P H 1996 *Phys. Plasmas* **3** 863
Medvedev M V, Shevchenko V I, Diamond P H and Galinsky V L 1997 *Phys. Plasmas* **4** 1257
Melville S, Schekochihin A A and Kunz M W 2016 *Mon. Not. R. Astron. Soc.* **459** 2701
Mikhailovskii A B and Tsypin V S 1971 *Plasma Phys.* **13** 785
Montgomery D 1983 *NASA Conf. Publication* vol 228
Ng C S and Bhattacharjee A 1996 *Astrophys. J.* **465** 845
Ofman L 2010 *Living Rev. Sol. Phys.* **7** 4
O'Neil T 1965 *Phys. Fluids* **8** 2255
Parker E N 1958 *Phys. Rev.* **109** 1874
Passot T, Sulem P L and Hunana P 2012 *Phys. Plasmas* **19** 082113
Quataert E 2001 *Probing the Physics of Active Galactic Nuclei by Multiwavelength Monitoring* (*Astronomical Society of the Pacific Conference Series* vol 224) ed B M Peterson *et al* (San Francisco, CA: Astronomical Society of the Pacific) p 71
Quataert E and Gruzinov A 1999 *Astrophys. J.* **520** 248
Ramos J J 2005 *Phys. Plasmas* **12** 052102
Rincon F, Schekochihin A A and Cowley S C 2015 *Mon. Not. R. Astron. Soc.* **447** L45
Riquelme M A, Quataert E and Verscharen D 2016 *Astrophys. J.* **824** 123
Rosenbluth M N 1956 Stability of the pinch *Technical Report* LA-2030 (Los Alamos, NM: Los Alamos Scientific Lab.)
Rosin M S, Schekochihin A A, Rincon F and Cowley S C 2011 *Mon. Not. R. Astron. Soc.* **413** 7
Santos-Lima R *et al* 2014 *Astrophys. J.* **781** 84
Schekochihin A A *et al* 2009 *Astrophys. J. Suppl.* **182** 310
Schekochihin A A, Cowley S C, Kulsrud R M, Rosin M S and Heinemann T 2008 *Phys. Rev. Lett.* **100** 081301
Schekochihin A A, Cowley S C, Rincon F and Rosin M S 2010 *Mon. Not. R. Astron. Soc.* **405** 291
Schekochihin A A, Nazarenko S V and Yousef T A 2012 *Phys. Rev.* E **85** 036406
Schlickeiser R 2015 *Phys. Plasmas* **22** 091502

Servidio S *et al* 2015 *J. Plasma Phys.* **81** 325810107
Sharma P, Hammett G W, Quataert E and Stone J M 2006 *Astrophys. J.* **637** 952
Sharma P, Quataert E, Hammett G W and Stone J M 2007 *Astrophys. J.* **667** 714
Snyder P B, Hammett G W and Dorland W 1997 *Phys. Plasmas* **4** 3974
Squire J and Quataert E 2017 in preparation
Squire J, Quataert E and Schekochihin A A 2016 *Astrophys. J. Lett.* **830** L25
Stoneham R J 1981 *Origin of Cosmic Rays* (*IAU Symp*. vol 94) ed G Setti p 255
Sulem P L and Passot T 2015 *J. Plasma Phys.* **81** 325810103
Tenerani A and Velli M 2013 *J. Geophys. Res.* **118** 7507
Verscharen D, Chandran B D G, Klein K G and Quataert E 2016 *Astrophys. J.* **831** 128
Völk H J and Cesarsky C J 1982 *Z. Naturforsch.* **37** 809
Yoon P H, Wu C S and de Assis A S 1993 *Phys. Fluids* B **5** 1971
Zhuravleva I *et al* 2014 *Nature* **515** 85